

This Page Is Inserted by IFW Operations  
and is not a part of the Official Record

## **BEST AVAILABLE IMAGES**

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

**IMAGES ARE BEST AVAILABLE COPY.**

As rescanning documents *will not* correct images,  
please do not report the images to the  
Image Problem Mailbox.

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 9, 2003, 23:15:47 ; Search time 32 Seconds  
(without alignments)  
1010.172 Million cell updates/sec

Title: US-09-818-247-1

Perfect score: 764  
Sequence: 1 MLFVLITCLAVFPAISTKS.....DFLLSSVAAEAQDGPQEA 764

Scoring table:

Gapop 60.0 , Gapext 60.0

Searched: 328717 seqs, 42310858 residues

Word size : 0

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database :

Issued Patents AA: \*  
1: /cgn2\_6/prodata/1/aa/5A\_COMB.pep: \*  
2: /cgn2\_6/prodata/1/aa/5B\_COMB.pep: \*  
3: /cgn2\_6/prodata/1/aa/5A\_COMB.pep: \*  
4: /cgn2\_6/prodata/1/aa/5B\_COMB.pep: \*  
5: /cgn2\_6/prodata/1/aa/PCITUS\_COMB.pep: \*  
6: /cgn2\_6/prodata/1/aa/backfile1.pep: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query	Match Length	ID	Description
1	746	97.6	746	US-08-434-000A-4	Sequence 4, Appl1
2	746	97.6	746	US-09-312-157-4	Sequence 4, Appl1
3	607	79.5	608	US-09-095-385-4	Sequence 4, Appl1
4	61	8.0	61	US-08-856-383-10	Sequence 10, Appl1
5	61	8.0	61	US-08-856-383-11	Sequence 11, Appl1
6	61	8.0	61	US-09-475-088-10	Sequence 10, Appl1
7	61	8.0	61	US-09-475-088-11	Sequence 11, Appl1
8	60	7.9	60	US-08-955-937A-6	Sequence 6, Appl1
9	60	7.9	60	US-09-300-985-6	Sequence 6, Appl1
10	53	5.9	109	US-08-961-564A-9	Sequence 9, Appl1
11	43	5.6	43	US-08-955-937A-11	Sequence 11, Appl1
12	43	5.6	43	US-09-300-985-11	Sequence 11, Appl1
13	31	4.1	31	US-08-856-383-4	Sequence 4, Appl1
14	29	3.8	31	US-09-475-088-4	Sequence 4, Appl1
15	29	3.8	46	US-08-955-937A-10	Sequence 10, Appl1
16	29	3.8	46	US-09-300-985-10	Sequence 10, Appl1
17	27	3.5	57	US-08-955-937A-5	Sequence 5, Appl1
18	27	3.5	57	US-09-300-985-5	Sequence 5, Appl1
19	21	2.7	757	US-08-434-000A-6	Sequence 6, Appl1
20	21	2.7	757	US-09-312-157-6	Sequence 6, Appl1
21	19	2.5	40	US-08-856-383-6	Sequence 6, Appl1
22	19	2.5	40	US-09-475-088-6	Sequence 6, Appl1
23	17	2.2	769	US-08-434-000A-10	Sequence 10, Appl1
24	17	2.2	769	US-09-312-157-10	Sequence 10, Appl1
25	16	2.1	16	US-08-782-480-45	Sequence 45, Appl1
26	16	2.1	16	US-08-954-211-45	Sequence 45, Appl1
27	16	2.1	16	US-09-005-167A-45	Sequence 45, Appl1

#### ALIGNMENTS

28	16	2.1	16	4	US-09-176-741B-45	Sequence 45, Appl1
29	16	2.1	771	3	US-08-434-000A-8	Sequence 8, Appl1
30	16	2.1	771	4	US-09-312-157-8	Sequence 8, Appl1
31	13	1.7	773	3	US-08-434-000A-2	Sequence 2, Appl1
32	13	1.7	773	4	US-09-312-157-2	Sequence 2, Appl1
33	12	1.6	16	2	US-08-656-906-1	Sequence 1, Appl1
34	12	1.6	16	3	US-09-217-847-1	Sequence 1, Appl1
35	9	1.2	624	2	US-08-642-406A-22	Sequence 22, Appl1
36	9	1.2	624	4	US-09-199-534-22	Sequence 22, Appl1
37	9	1.2	624	3	US-09-199-534-22	Sequence 22, Appl1
38	8	1.0	60	3	US-08-961-564A-6	Sequence 6, Appl1
39	8	1.0	417	4	US-09-252-991A-27213	Sequence 27213, A
40	8	1.0	777	2	US-08-674-678-3	Sequence 3, Appl1
41	8	1.0	777	3	US-08-643-839-3	Sequence 3, Appl1
42	8	1.0	777	4	US-09-348-886-3	Sequence 3, Appl1
43	8	1.0	802	4	US-09-173-151A-33	Sequence 33, Appl1
44	8	1.0	1298	1	US-08-222-616-33	Sequence 33, Appl1
45	8	1.0	1298	1	US-08-340-011-2	Sequence 2, Appl1

RESULT 1  
US-08-434-000A-4  
Sequence 4, Application US/08434000A  
Patent No. 6046037  
GENERAL INFORMATION:  
APPLICANT: ANDREW C. HIATT, JULIAN  
TITLE OF INVENTION: IMMUNOGLOBULINS CONTAINING PROTECTION  
TITLE OF INVENTION: PROTEINS IN PLANTS AND THEIR USES  
NUMBER OF SEQUENCES: 19  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Lyon & Lyon  
STREET: 633 West Fifth Street  
STREET: Suite 4700  
CITY: Los Angeles  
STATE: California  
COUNTRY: U.S.A.  
ZIP: 90071  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb  
MEDIUM TYPE: storage  
COMPUTER: IBM compatible  
OPERATING SYSTEM: IBM P.C. DOS 5.0  
SOFTWARE: Word Perfect 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/434,000A  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
PRIOR APPLICATION DATA: including application  
PRIOR APPLICATION DATA: described below:  
APPLICATION NUMBER: 08/367,395  
FILING DATE: 12/30/94  
ATTORNEY/AGENT INFORMATION:  
NAME: Guise, Jeffrey W.  
REGISTRATION NUMBER: 34,613  
REFERENCE/DOCKET NUMBER: 212/127  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (619) 552-8400  
TELEFAX: (619) 552-0159  
TELEX: 67-3510  
TELEX: SEQUENCE LISTING  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 746 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
DESCRIPTION: Human Polymunoglobulin Receptor  
US-08-434-000A-4

Query Match 97.6%; Score 746; DB 3; Length 746;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 746; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

19 KSPIGPEEVNVEGNSVITCYPTSYNRRHTRKRYWCRGARGCITLISSEGVSSKY 78  
 1 KSPIGPEEVNVEGNSVITCYPTSYNRRHTRKRYWCRGARGCITLISSEGVSSKY 60

79 AGRALNTPENGTFVNNIAQLSODDSGRYKCGCLINSRGLSFDVSLSEVSGPGLNDTK 138  
 61 AGRALNTPENGTFVNNIAQLSODDSGRYKCGCLINSRGLSFDVSLSEVSGPGLNDTK 120

139 VTTVDLGRVTINCPFKTEENAKRKSLYKQIGLYPVLYDSSGYVNPNTGRIRLDIOGT 198  
 121 VTTVDLGRVTINCPFKTEENAKRKSLYKQIGLYPVLYDSSGYVNPNTGRIRLDIOGT 180

199 GOLFSVYINQLRLSDAGQYLCQAGDSDNSNKNADLOYLKPPELVEDLGSYTFPCA 258  
 181 GOLFSVYINQLRLSDAGQYLCQAGDSDNSNKNADLOYLKPPELVEDLGSYTFPCA 240

259 LGPEVANNVAKFLCROSSGGENCDVYNTLCKRAPAFEGRIILNPODKGSFSVYITGLRKE 318  
 241 LGPEVANNVAKFLCROSSGGENCDVYNTLCKRAPAFEGRIILNPODKGSFSVYITGLRKE 300

319 DAGRYLCGAHSDGQLOEGSPIOAMOLFVNEESTIPRSPYVGVAGSSVAVLCPYNNRKS 378  
 301 DAGRYLCGAHSDGQLOEGSPIOAMOLFVNEESTIPRSPYVGVAGSSVAVLCPYNNRKS 360

379 KSIIKWCMEGAONGRCPLVDSEGVNKAQYGRSLLEPENGFTVILNOLTSRDAGF 438  
 361 KSIIKWCMEGAONGRCPLVDSEGVNKAQYGRSLLEPENGFTVILNOLTSRDAGF 420

439 YWCLTNGDPLMRTVEIKIIEGEPNLKVGNTAVLGETLKYPCHPCKFSSEYKWCWK 498  
 421 YWCLTNGDPLMRTVEIKIIEGEPNLKVGNTAVLGETLKYPCHPCKFSSEYKWCWK 480

499 NNTGCOALPSODEGSKAFVNCDENSRLVSLTLNLTADDEGMWGVQGHFYGETAAV 558  
 481 NNTGCOALPSODEGSKAFVNCDENSRLVSLTLNLTADDEGMWGVQGHFYGETAAV 540

559 YVAVEERKAGSRDVSILAKADAPDEKVLDSGFREIENKAIDOPRLFAEKKVADTRDA 618  
 541 YVAVEERKAGSRDVSILAKADAPDEKVLDSGFREIENKAIDOPRLFAEKKVADTRDA 600

619 DCSRASVDSSSEEDGSSRALVSTLPLGLVLAAGAVALGVARARHRKNVDVRSIRSYR 678  
 601 DCSRASVDSSSEEDGSSRALVSTLPLGLVLAAGAVALGVARARHRKNVDVRSIRSYR 660

679 TDISMSDFENSREFGANDMGASSITQETSLGKKEFFVATTESTTETKEPKAKKSSKEE 738  
 661 TDISMSDFENSREFGANDMGASSITQETSLGKKEFFVATTESTTETKEPKAKKSSKEE 720

739 AEMAYKDFLLOSTVAEAADGPOEA 764  
 721 AEMAYKDFLLOSTVAEAADGPOEA 746

STATE: California  
 COUNTRY: U.S.A.  
 ZIP: 90071  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: 3.5" Diskette, 1.44 Mb  
 storage  
 COMPUTER: IBM Compatible  
 OPERATING SYSTEM: IBM P.C. DOS 5.0  
 SOFTWARE: Word Perfect 5.1  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/09/312,157  
 FILING DATE: 14-May-1999  
 CLASSIFICATION: <unknown>  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: 08/434,000  
 FILING DATE: <unknown>  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Guise, Jeffrey W.  
 REGISTRATION NUMBER: 34,613  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (619) 552-8400  
 TELEFAX: (619) 552-0159  
 TELEX: 67-351  
 SEQUENCE LISTING  
 INFORMATION FOR SEQ ID NO: 4:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 746 amino acids  
 TYPE: amino acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 DESCRIPTION: Human Polymunoglobulin Receptor  
 SEQUENCE DESCRIPTION: SEQ ID NO: 4

US-09-312-157-4

Query Match 97.6%; Score 746; DB 4; Length 746;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 746; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

19 KSPIGPEEVNVEGNSVITCYPTSYNRRHTRKRYWCRGARGCITLISSEGVSSKY 78  
 1 KSPIGPEEVNVEGNSVITCYPTSYNRRHTRKRYWCRGARGCITLISSEGVSSKY 60

79 AGRALNTPENGTFVNNIAQLSODDSGRYKCGCLINSRGLSFDVSLSEVSGPGLNDTK 138  
 61 AGRALNTPENGTFVNNIAQLSODDSGRYKCGCLINSRGLSFDVSLSEVSGPGLNDTK 120

139 VTTVDLGRVTINCPFKTEENAKRKSLYKQIGLYPVLYDSSGYVNPNTGRIRLDIOGT 198  
 121 VTTVDLGRVTINCPFKTEENAKRKSLYKQIGLYPVLYDSSGYVNPNTGRIRLDIOGT 180

199 GOLFSVYINQLRLSDAGQYLCQAGDSDNSNKNADLOYLKPPELVEDLGSYTFPCA 258  
 181 GOLFSVYINQLRLSDAGQYLCQAGDSDNSNKNADLOYLKPPELVEDLGSYTFPCA 240

259 LGPEVANNVAKFLCROSSGGENCDVYNTLCKRAPAFEGRIILNPODKGSFSVYITGLRKE 318  
 241 LGPEVANNVAKFLCROSSGGENCDVYNTLCKRAPAFEGRIILNPODKGSFSVYITGLRKE 300

319 DAGRYLCGAHSDGQLOEGSPIOAMOLFVNEESTIPRSPYVGVAGSSVAVLCPYNNRKS 378  
 301 DAGRYLCGAHSDGQLOEGSPIOAMOLFVNEESTIPRSPYVGVAGSSVAVLCPYNNRKS 360

379 KSIIKWCMEGAONGRCPLVDSEGVNKAQYGRSLLEPENGFTVILNOLTSRDAGF 438  
 361 KSIIKWCMEGAONGRCPLVDSEGVNKAQYGRSLLEPENGFTVILNOLTSRDAGF 420

439 YWCLTNGDPLMRTVEIKIIEGEPNLKVGNTAVLGETLKYPCHPCKFSSEYKWCWK 498  
 421 YWCLTNGDPLMRTVEIKIIEGEPNLKVGNTAVLGETLKYPCHPCKFSSEYKWCWK 480

499 NNTGCOALPSODEGSKAFVNCDENSRLVSLTLNLTADDEGMWGVQGHFYGETAAV 558  
 421 NNTGCOALPSODEGSKAFVNCDENSRLVSLTLNLTADDEGMWGVQGHFYGETAAV 558

RESULT 2  
 US-09-312-157-4  
 : Sequence 4, Application US/09312157  
 : Patent No. 6303341  
 : GENERAL INFORMATION:  
 : APPLICANT: ANDREW C. HIATT, JULIAN  
 : K.-C. MA, THOMAS LEHNER  
 : TITLE OF INVENTION: IMMUNOGLOBULINS CONTAINING PROTECTION  
 : PROTEINS IN PLANTS AND THEIR USES  
 : NUMBER OF SEQUENCES: 19  
 : CORRESPONDENCE ADDRESS:  
 : ADDRESS: Lyon & Lyon  
 : STREET: 633 West Fifth Street  
 : Suite 4700  
 : CITY: Los Angeles

Db 481 NNTGCOALPSODEPESKAFVNCDENSRLVSLTLNLTTRADEGWTWCYKQGHFYGETAAV 540

Qy 559 YVAEERKAAGSRDVSIAKADAPDEKVLDSGFREIENKAIQDPRLEFAEERKAVADTRDA 618

Db 541 YVAEERKAAGSRDVSIAKADAPDEKVLDSGFREIENKAIQDPRLEFAEERKAVADTRDA 600

Qy 619 DGSNASVDSGSSSEOGGSSRALVSTLPLGLVAVGAVGAVARHRKKNVDVRSISYR 678

Db 601 DGSNASVDSGSSSEOGGSSRALVSTLPLGLVAVGAVGAVARHRKKNVDVRSISYR 660

Qy 679 TDISMSPFENSREGANDNMGASSITQETSLGKKEEFPATTESTETETKEPKAKRSSKEE 738

Db 661 TDISMSPFENSREGANDNMGASSITQETSLGKKEEFPATTESTETETKEPKAKRSSKEE 720

Qy 739 AEMAYKDFLLQSTVAEAADGPOEA 764

Db 721 AEMAYKDFLLQSTVAEAADGPOEA 746

# RESULT 3

US-09-095-385-4

Sequence 4, Application US/09095385

Patent No. 6300104

## GENERAL INFORMATION:

APPLICANT: Morrison, Sherie L.

TITLE OF INVENTION: SECRETORY IMMUNOGLOBULIN PRODUCED BY SINGLE CELLS AND METHODS FOR MAKING AND USING

TITLE OF INVENTION: SAME

NUMBER OF SEQUENCES: 4

CORRESPONDENCE ADDRESS:

ADDRESSEE: Merchant, Gould, Smith, Edell, Welter & Schmidt

STREET: 11150 Santa Monica Boulevard, Suite 400

CITY: Los Angeles

STATE: CA

COUNTRY: USA

ZIP: 90025

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

SOFTWARE: FASTSEQ for Windows Version 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/095.385

FILING DATE: 09-JUN-1998

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 60/050,969

FILING DATE: 19-JUN-1997

ATTORNEY/AGENT INFORMATION:

NAME: Canady, Karen S

REGISTRATION NUMBER: 39,927

REFERENCE/DOCKET NUMBER: 30435.45USU1

TELECOMMUNICATION INFORMATION:

TELEPHONE: 310 445-1140

TELEFAX: 310 445-9031

TELEX:

INFORMATION FOR SEQ ID NO: 4:

SEQUENCE CHARACTERISTICS:

LENGTH: 608 amino acids

TYPE: amino acid

STRANDEDNESS: unknown

TOPOLOGY: linear

MOLECULE TYPE: protein

US-09-095-385-4

Query Match 79.5%; Score 607; DB 4; Length 608;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 607; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MLFVLTCLAVPAISTKSPFGPEEVNSVEGNSVITCYPTSVNRHTRKRWCKOGA 60

Db 1 MLFVLTCLAVPAISTKSPFGPEEVNSVEGNSVITCYPTSVNRHTRKRWCKOGA 60

Qy 61 RGCITLISSEGVSSYKAGRANLTNPENGTFFVNIATQSDDSGRYKCGLGINSGLS 120

Db 61 RGCITLISSEGVSSYKAGRANLTNPENGTFFVNIATQSDDSGRYKCGLGINSGLS 120

Qy 121 FDSVLEVSQSGPGLNDTKVTVLDGRVTTINCPKTEMAOKRSLYQIGLYPLVYDSS 180

Db 121 FDSVLEVSQSGPGLNDTKVTVLDGRVTTINCPKTEMAOKRSLYQIGLYPLVYDSS 180

Qy 181 GYVNPNTGRIRLDIOGTGOLLFSVNIQRLSPAGQYLCQAGDSSNNKNADLYLKP 240

Db 181 GYVNPNTGRIRLDIOGTGOLLFSVNIQRLSPAGQYLCQAGDSSNNKNADLYLKP 240

Qy 241 EPELVYEDLGSVTFHCALGPEVANVAKFLCROSSGECNVVNTLGRAPAEGRILLN 300

Db 241 EPELVYEDLGSVTFHCALGPEVANVAKFLCROSSGECNVVNTLGRAPAEGRILLN 300

Qy 301 PODKGSFVSITGLRKEDAGRYLCGASHDQLOEGSPIDAMOLFVNEESTIPSPVVK 360

Db 301 PODKGSFVSITGLRKEDAGRYLCGASHDQLOEGSPIDAMOLFVNEESTIPSPVVK 360

Qy 361 GVAGSSVAVLCPYNRKESKSIKYWCLMEGAQNGRCPLLVDSSEGVAKQYEGRLSLEEPG 420

Db 361 GVAGSSVAVLCPYNRKESKSIKYWCLMEGAQNGRCPLLVDSSEGVAKQYEGRLSLEEPG 420

Qy 421 NGFTVILNOLTSRDAGFYWCLNTGDTLMTTVEIKIIEBPNLKVGNVAVLGEFLKY 480

Db 421 NGFTVILNOLTSRDAGFYWCLNTGDTLMTTVEIKIIEBPNLKVGNVAVLGEFLKY 480

Qy 481 PCHEPKFESSYKXWCMWNTGCOALPSODEGPKAVVNDENSRLVSLTLNLTTRADEG 540

Db 481 PCHEPKFESSYKXWCMWNTGCOALPSODEGPKAVVNDENSRLVSLTLNLTTRADEG 540

Qy 541 WYMGVYKQGHFYGETAAVYVAEERKAAGSRDVSIAKADAPDEKVLDSGFREIENKAIQ 600

Db 541 WYMGVYKQGHFYGETAAVYVAEERKAAGSRDVSIAKADAPDEKVLDSGFREIENKAIQ 600

Qy 601 DPRLEAE 607

Db 601 DPRLEAE 607

# RESULT 4

US-08-856-383-10

Sequence 10, Application US/08856383

Patent No. 6042833

## GENERAL INFORMATION:

APPLICANT: Mostov, Keith E.

TITLE OF INVENTION: Cellular Internalization of Pig R Stalk

TITLE OF INVENTION: and Associated Ligands

NUMBER OF SEQUENCES: 11

CORRESPONDENCE ADDRESS:

ADDRESSEE: Townsend and Townsend and Crew LLP

STREET: Two Embarcadero Center, Eighth Floor

CITY: San Francisco

STATE: California

COUNTRY: USA

ZIP: 94111-3834

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/856.383

FILING DATE: No. 6042833 yet assigned

CLASSIFICATION: 530

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 60/018,958

FILING DATE: 04-JUN-1996

ATTORNEY/AGENT INFORMATION:

NAME: Weber, Kenneth A.

REGISTRATION NUMBER: 31,677  
REFERENCE/DOCKET NUMBER: 02307E-067910PC  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 576-0200  
TELEFAX: (415) 576-0300  
INFORMATION FOR SEQ ID NO: 10:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 61 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-856-383-10

Query Match 8.0%; Score 61; DB 3; Length 61;  
Best Local Similarity 100.0%; Pred. No. 1.5e-54;  
Matches 61; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

578 ADAAPDEKVLDSGFREIENKAIDPRLFAEKAIVADTRDQADGSRASVDSGSSSEEGCGSS 637  
|||||  
1 ADAAPDEKVLDSGFREIENKAIDPRLFAEKAIVADTRDQADGSRASVDSGSSSEEGCGSS 60

Oy 638 R 638  
Db 61 R 61

RESULT 5  
US-08-856-383-11  
Sequence 11, Application US/08856383  
Patent No. 6042833  
GENERAL INFORMATION:  
APPLICANT: Mostov, Keith E.  
TITLE OF INVENTION: Cellular Internalization of p19R Stalk  
NUMBER OF SEQUENCES: 11  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Townsend and Townsend and Crew LLP  
STREET: Two Embarcadero Center, Eighth Floor  
CITY: San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94111-3834  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/856,383  
FILING DATE: No. 6042833 yet assigned  
CLASSIFICATION: 530  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 60/018,958  
FILING DATE: 04-JUN-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Weber, Kenneth A.  
REGISTRATION NUMBER: 31,677  
REFERENCE/DOCKET NUMBER: 02307E-067910PC  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 576-0200  
TELEFAX: (415) 576-0300  
INFORMATION FOR SEQ ID NO: 11:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 61 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-856-383-11

Query Match 8.0%; Score 61; DB 3; Length 61;

Best Local Similarity 100.0%; Pred. No. 1.5e-54;  
Matches 61; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 578 ADAAPDEKVLDSGFREIENKAIDPRLFAEKAIVADTRDQADGSRASVDSGSSSEEGCGSS 637  
|||||  
Db 1 ADAAPDEKVLDSGFREIENKAIDPRLFAEKAIVADTRDQADGSRASVDSGSSSEEGCGSS 60

Oy 638 R 638  
Db 61 R 61

RESULT 6  
US-09-475-088-10  
Sequence 10, Application US/09475088  
Patent No. 6340743  
GENERAL INFORMATION:  
APPLICANT: Mostov, Keith E.  
TITLE OF INVENTION: Cellular Internalization of p19R Stalk  
NUMBER OF SEQUENCES: 11  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Townsend and Townsend and Crew LLP  
STREET: Two Embarcadero Center, Eighth Floor  
CITY: San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94111-3834  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/475,088  
FILING DATE: 30-Dec-1999  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/856,383  
FILING DATE: <Unknown>  
ATTORNEY/AGENT INFORMATION:  
NAME: Weber, Kenneth A.  
REGISTRATION NUMBER: 31,677  
REFERENCE/DOCKET NUMBER: 02307E-067910PC  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 576-0200  
TELEFAX: (415) 576-0300  
INFORMATION FOR SEQ ID NO: 10:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 61 amino acids  
TYPE: amino acid  
STRANDEDNESS: <Unknown>  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
SEQUENCE DESCRIPTION: SEQ ID NO: 10:  
US-09-475-088-10

Query Match 8.0%; Score 61; DB 4; Length 61;  
Best Local Similarity 100.0%; Pred. No. 1.5e-54;  
Matches 61; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 578 ADAAPDEKVLDSGFREIENKAIDPRLFAEKAIVADTRDQADGSRASVDSGSSSEEGCGSS 637  
|||||  
Db 1 ADAAPDEKVLDSGFREIENKAIDPRLFAEKAIVADTRDQADGSRASVDSGSSSEEGCGSS 60

Oy 638 R 638  
Db 61 R 61

RESULT 7  
US-09-475-088-11

Sequence 11, Application US/09475088  
Patent No. 6340743  
GENERAL INFORMATION:  
APPLICANT: Mostov, Keith E.  
TITLE OF INVENTION: Richman-Eisenstat, Janice  
and Associated Ligands  
CELLULAR INTERNALIZATION OF PIGR STALK  
NUMBER OF SEQUENCES: 11  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Townsend and Townsend and Crew LLP  
CITY: San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94111-3834  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/475,088  
FILING DATE: 30-Dec-1999  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/856,383  
FILING DATE: <Unknown>  
ATTORNEY/AGENT INFORMATION:  
NAME: Weber, Kenneth A.  
REGISTRATION NUMBER: 31,677  
REFERENCE/DOCKET NUMBER: 02307E-067910PC  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 576-0200  
TELEFAX: (415) 576-0300  
INFORMATION FOR SEQ ID NO: 11:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 61 amino acids  
TYPE: amino acid  
STRANDEDNESS: <Unknown>  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
SEQUENCE DESCRIPTION: SEQ ID NO: 11:  
US-09-475-088-11

Query Match  
Best Local Similarity 8.0%; Score 61; DB 4; Length 61;  
Matches 61; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 578 ADAPDEKVLDSGFREIENKAIQDPRLEAEKAVADTRDQADGSRASVDSGSGSEGGSS 637  
1 ADAPDEKVLDSGFREIENKAIQDPRLEAEKAVADTRDQADGSRASVDSGSGSEGGSS 60

QY 638 R 638  
DB 61 R 61

RESULT 8  
US-08-955-937A-6  
Sequence 6, Application US/08955937A  
Patent No. 6020161  
GENERAL INFORMATION:  
APPLICANT: WU, SHUJIAN  
APPLICANT: SWEET, RAYMOND  
APPLICANT: TRUNEH, ALEMGEGED  
TITLE OF INVENTION: PIGR-1, A MEMBER OF IMMUNOGLOBULIN  
GENE SUPERFAMILY  
NUMBER OF SEQUENCES: 14  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: RATNER & PRESTIA  
STREET: P. O. BOX 980  
CITY: VALLEY FORGE  
STATE: PA

COUNTRY: USA  
ZIP: 19482  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FASTSEQ for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/955,937A  
FILING DATE: 17-OCT-1997  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 60/056,152  
FILING DATE: 19-AUG-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: PRESTIA, PAUL F.  
REGISTRATION NUMBER: 23,031  
REFERENCE/DOCKET NUMBER: GH-70228  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 610-407-0700  
TELEFAX: 610-407-0701  
TELEX: 846169  
INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 60 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-955-937A-6

Query Match  
Best Local Similarity 7.9%; Score 60; DB 3; Length 60;  
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 353 PRSPVYVKGAVGSSVAVALCPYNRKESKSIKYWCLMEGAONGRCPLLVDSEGVYKAYEGR 412  
1 PRSPVYVKGAVGSSVAVALCPYNRKESKSIKYWCLMEGAONGRCPLLVDSEGVYKAYEGR 60

DB

RESULT 9  
US-09-300-985-6  
Sequence 6, Application US/09300985A  
Patent No. 6232441  
GENERAL INFORMATION:  
APPLICANT: WU, SHUJIAN  
APPLICANT: SWEET, RAYMOND W.  
APPLICANT: TRUNEH, ALEMGEGED  
TITLE OF INVENTION: PIGR-1, A MEMBER OF IMMUNOGLOBULIN GENE  
FILE REFERENCE: GH-70228-1  
CURRENT APPLICATION NUMBER: US/09/300,985A  
CURRENT FILING DATE: 1999-04-28  
EARLIER APPLICATION NUMBER: US 08/955,937  
EARLIER FILING DATE: 1997-10-22  
EARLIER APPLICATION NUMBER: US 60/056,152  
EARLIER FILING DATE: 1997-08-19  
NUMBER OF SEQ ID NOS: 14  
SOFTWARE: FASTSEQ for Windows Version 3.0  
SEQ ID NO: 6  
LENGTH: 60  
TYPE: PRT  
ORGANISM: HOMO SAPIENS  
US-09-300-985-6

Query Match  
Best Local Similarity 7.9%; Score 60; DB 3; Length 60;  
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 353 PRSPVYVKGAVGSSVAVALCPYNRKESKSIKYWCLMEGAONGRCPLLVDSEGVYKAYEGR 412  
1 PRSPVYVKGAVGSSVAVALCPYNRKESKSIKYWCLMEGAONGRCPLLVDSEGVYKAYEGR 60

DB

RESULT 10  
US-08-961-564A-9  
Sequence 9, Application US/08961564A  
Patent No. 6114515  
GENERAL INFORMATION:  
APPLICANT: WU, SHUJIAN  
APPLICANT: SWEET, RAYMOND  
APPLICANT: TRUNEH, ALEMSEGED  
TITLE OF INVENTION: PIGR-1, A MEMBER OF IMMUNOGLOBULIN  
TITLE OF INVENTION: GENE SUPERFAMILY  
NUMBER OF SEQUENCES: 9  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: RATNER & PRESTIA  
STREET: P.O. BOX 980  
CITY: VALLEY FORGE  
STATE: PA  
COUNTRY: USA  
ZIP: 19482  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FASTSEQ for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/961,564A  
FILING DATE: 30-OCT-1997  
CLASSIFICATION: 536  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 60/056,935  
FILING DATE: 25-AUG-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: PRESTIA, PAUL F  
REGISTRATION NUMBER: 23,031  
REFERENCE/DOCKET NUMBER: GH-70236  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 610-407-0700  
TELEFAX: 610-407-0701  
TELEX: 846169  
INFORMATION FOR SEQ ID NO: 9:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 109 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-961-564A-9  
Query Match 6.9%; Score 53; DB 3; Length 109;  
Best Local Similarity 100.0%; Pred. No. 3.9e-46;  
Matches 53; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Oy 64 CITLISSEGVSKYAGRAVLTFNPENGFVVIQAQLSDSDSGRYCGLGINS 116  
Db 45 CITLISSEGVSKYAGRAVLTFNPENGFVVIQAQLSDSDSGRYCGLGINS 97  
RESULT 11  
US-08-955-937A-11  
Sequence 11, Application US/08955937A  
Patent No. 6020161  
GENERAL INFORMATION:  
APPLICANT: WU, SHUJIAN  
APPLICANT: SWEET, RAYMOND  
APPLICANT: TRUNEH, ALEMSEGED  
TITLE OF INVENTION: PIGR-1, A MEMBER OF IMMUNOGLOBULIN  
TITLE OF INVENTION: GENE SUPERFAMILY  
NUMBER OF SEQUENCES: 14  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: RATNER & PRESTIA  
STREET: P.O. BOX 980  
CITY: VALLEY FORGE

STATE: PA  
COUNTRY: USA  
ZIP: 19482  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FASTSEQ for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/955,937A  
FILING DATE: 17-OCT-1997  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 60/056,152  
FILING DATE: 19-AUG-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: PRESTIA, PAUL F  
REGISTRATION NUMBER: 23,031  
REFERENCE/DOCKET NUMBER: GH-70228  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 610-407-0700  
TELEFAX: 610-407-0701  
TELEX: 846169  
INFORMATION FOR SEQ ID NO: 11:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 43 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-955-937A-11  
Query Match 5.6%; Score 43; DB 3; Length 43;  
Best Local Similarity 100.0%; Pred. No. 2.8e-36;  
Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Oy 413 LSLLEPENGTFVILNQLTSRDAGFYWCITNGDTLMTPTVEI 455  
Db 1 LSLLEPENGTFVILNQLTSRDAGFYWCITNGDTLMTPTVEI 43  
RESULT 12  
US-09-300-985-11  
Sequence 11, Application US/09300985A  
Patent No. 6232441  
GENERAL INFORMATION:  
APPLICANT: WU, SHUJIAN  
APPLICANT: SWEET, RAYMOND W.  
APPLICANT: TRUNEH, ALEMSEGED  
APPLICANT: HUBLE, MARK ROBERT  
TITLE OF INVENTION: PIGR-1, A MEMBER OF IMMUNOGLOBULIN GENE  
FILE REFERENCE: GH-70228-1  
CURRENT APPLICATION NUMBER: US/09/300,985A  
CURRENT FILING DATE: 1999-04-28  
EARLIER APPLICATION NUMBER: US 08/955,937  
EARLIER FILING DATE: 1997-10-22  
EARLIER APPLICATION NUMBER: US 60/056,152  
EARLIER FILING DATE: 1997-08-19  
NUMBER OF SEQ ID NOS: 14  
SOFTWARE: FASTSEQ for Windows Version 3.0  
SEQ ID NO 11  
LENGTH: 43  
TYPE: PRT  
ORGANISM: HOMO SAPIENS  
US-09-300-985-11  
Query Match 5.6%; Score 43; DB 3; Length 43;  
Best Local Similarity 100.0%; Pred. No. 2.8e-36;  
Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Oy 413 LSLLEPENGTFVILNQLTSRDAGFYWCITNGDTLMTPTVEI 455  
Db 1 LSLLEPENGTFVILNQLTSRDAGFYWCITNGDTLMTPTVEI 43

Db 1 LSLLEPONGTFTVILNOLTSRDAGFTWCLNGDTLMTTVEI 43

## RESULT 13

US-08-856-383-4

Sequence 4, Application US/08856383

Patent No. 6042833

GENERAL INFORMATION:

APPLICANT: Mostov, Keith E.

APPLICANT: Richman-Eisenstat, Janice

TITLE OF INVENTION: Cellular Internalization of p19r stalk

TITLE OF INVENTION: and Associated Ligands

NUMBER OF SEQUENCES: 11

CORRESPONDENCE ADDRESS:

ADDRESSEE: Townsend and Townsend and Crew LLP

STREET: Two Embarcadero Center, Eighth Floor

CITY: San Francisco

STATE: California

COUNTRY: USA

ZIP: 94111-3834

COMPUTER READABLE FORM:

MEDIUM TYPE: floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/856,383

FILING DATE: No. 6042833 yet assigned

CLASSIFICATION: 530

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 60/018,958

FILING DATE: 04-JUN-1996

ATTORNEY/AGENT INFORMATION:

NAME: Weber, Kenneth A.

REGISTRATION NUMBER: 31,677

REFERENCE/DOCKET NUMBER: 02307E-067910PC

TELECOMMUNICATION INFORMATION:

TELEPHONE: (415) 576-0200

TELEFAX: (415) 576-0300

INFORMATION FOR SEQ ID NO: 4:

SEQUENCE CHARACTERISTICS:

LENGTH: 31 amino acids

TYPE: amino acid

STRANDEDNESS:

TOPOLOGY: linear

MOLECULE TYPE: peptide

US-08-856-383-4

Query Match 4.18; Score 31; DB 3; Length 31;

Best Local Similarity 100.0%; Pred. No. 4e-24;

Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 608 EKAVADTRDOADGSRASVDSGSSSEOGGSSR 638

Db 1 EKAVADTRDOADGSRASVDSGSSSEOGGSSR 31

## RESULT 14

US-09-475-088-4

Sequence 4, Application US/09475088

Patent No. 6340743

GENERAL INFORMATION:

APPLICANT: Mostov, Keith E.

APPLICANT: Richman-Eisenstat, Janice

TITLE OF INVENTION: Cellular Internalization of p19r stalk

TITLE OF INVENTION: and Associated Ligands

NUMBER OF SEQUENCES: 11

CORRESPONDENCE ADDRESS:

ADDRESSEE: Townsend and Townsend and Crew LLP

STREET: Two Embarcadero Center, Eighth Floor

CITY: San Francisco

STATE: California

COUNTRY: USA

ZIP: 94111-3834

COMPUTER READABLE FORM:

MEDIUM TYPE: floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/475,088

FILING DATE: 30-Dec-1999

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/856,383

FILING DATE: <Unknown>

ATTORNEY/AGENT INFORMATION:

NAME: Weber, Kenneth A.

REGISTRATION NUMBER: 31,677

REFERENCE/DOCKET NUMBER: 02307E-067910PC

TELECOMMUNICATION INFORMATION:

TELEPHONE: (415) 576-0200

TELEFAX: (415) 576-0300

INFORMATION FOR SEQ ID NO: 4:

SEQUENCE CHARACTERISTICS:

LENGTH: 31 amino acids

TYPE: amino acid

STRANDEDNESS: <Unknown>

TOPOLOGY: linear

MOLECULE TYPE: peptide

SEQUENCE DESCRIPTION: SEQ ID NO: 4:

US-09-475-088-4

Query Match 4.18; Score 31; DB 4; Length 31;

Best Local Similarity 100.0%; Pred. No. 4e-24;

Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 608 EKAVADTRDOADGSRASVDSGSSSEOGGSSR 638

Db 1 EKAVADTRDOADGSRASVDSGSSSEOGGSSR 31

## RESULT 15

US-08-955-937A-10

Sequence 10, Application US/08955937A

Patent No. 6020161

GENERAL INFORMATION:

APPLICANT: WU, SHUJIAN

APPLICANT: SWEET, RAYMOND

APPLICANT: TRUNH, ALEMSSEG

TITLE OF INVENTION: p19r-1, A MEMBER OF IMMUNOGLOBULIN

TITLE OF INVENTION: GENE SUPERFAMILY

NUMBER OF SEQUENCES: 14

CORRESPONDENCE ADDRESS:

ADDRESSEE: RATNER & PRESTIA

STREET: P.O. BOX 980

CITY: VALLEY FORGE

STATE: PA

COUNTRY: USA

ZIP: 19482

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

SOFTWARE: FastSeq for Windows Version 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/955,937A

FILING DATE: 17-OCT-1997

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 60/056,152

FILING DATE: 19-AUG-1997

ATTORNEY/AGENT INFORMATION:

NAME: PRESTIA, PAUL F

REGISTRATION NUMBER: 23,031

REFERENCE/DOCKET NUMBER: GH-70228



TELECOMMUNICATION INFORMATION:  
TELEPHONE: 610-407-0700  
TELEFAX: 610-407-0701  
TELEX: 846169  
INFORMATION FOR SEQ ID NO: 10:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 46 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-955-937A-10

Query Match 3.88; Score 29; DB 3; Length 46;  
Best Local Similarity 100.0%; Pred. No. 6.5e-22;  
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

7 99 QLSQDDSGRYKCGIGINSRGLSPDVSLEY 127  
|||  
18 QLSQDDSGRYKCGIGINSRGLSPDVSLEY 46

Search completed: September 9, 2003, 23:21:17  
Job time : 33 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 9, 2003, 23:18:43; Search time 65 Seconds

(without alignments)  
1613.846 Million cell updates/sec

Title: US-09-818-247-1

Perfect score: 764  
Sequence: 1 MLFLVTLCLAVFAIPSTKTS.....DFLLQSTVAEAADQGPQEA 764

Scoring table:

Gapop 60.0, Gapext 60.0

Searched: 513375 seqs, 137303645 residues

Word size: 0

Total number of hits satisfying chosen parameters: 513375

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database:

Published Applications\_AA:\*

- 1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*
- 2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep.\*
- 3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep.\*
- 4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep.\*
- 5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep.\*
- 6: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pep.\*
- 7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep.\*
- 8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*
- 9: /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.pep.\*
- 10: /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.pep.\*
- 11: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pep.\*
- 12: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep.\*
- 13: /cgn2\_6/ptodata/1/pubpaa/US10A\_PUBCOMB.pep.\*
- 14: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pep.\*
- 15: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep.\*
- 16: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep.\*
- 17: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep.\*
- 18: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	764	100.0	764	10	US-09-818-247-1 Sequence 1, Appl
2	764	100.0	764	10	US-09-881-353-59 Sequence 59, Appl
3	764	100.0	764	10	US-09-989-919-124 Sequence 124, Appl
4	746	97.6	746	10	US-09-982-107-4 Sequence 4, Appl
5	607	79.5	608	10	US-09-950-294-4 Sequence 4, Appl
6	243	31.8	243	12	US-09-969-748C-108 Sequence 108, Appl
7	205	26.8	602	14	US-10-047-542-51 Sequence 51, Appl
8	95	12.4	95	10	US-09-818-247-20 Sequence 20, Appl
9	94	12.3	102	12	US-09-969-748C-104 Sequence 104, Appl
10	90	11.8	90	10	US-09-818-247-17 Sequence 17, Appl
11	50	6.5	70	9	US-09-925-299-1218 Sequence 1218, Ap
12	50	6.5	70	11	US-09-925-299-1218 Sequence 1218, Ap
13	47	6.2	244	12	US-09-969-748C-109 Sequence 109, Appl
14	31	4.1	31	10	US-09-818-247-19 Sequence 19, Appl
15	22	2.9	94	12	US-09-969-748C-107 Sequence 107, Appl

16	21	2.7	757	10	US-09-818-247-2	Sequence 2, Appl
17	21	2.7	757	10	US-09-982-107-6	Sequence 6, Appl
18	20	2.6	102	12	US-09-969-748C-103	Sequence 103, Appl
19	18	2.4	18	12	US-09-969-748C-71	Sequence 71, Appl
20	18	2.4	23	12	US-09-969-748C-72	Sequence 72, Appl
21	18	2.4	24	12	US-09-969-748C-73	Sequence 73, Appl
22	17	2.2	249	12	US-09-969-748C-110	Sequence 110, Appl
23	17	2.2	769	10	US-09-818-247-3	Sequence 10, Appl
24	17	2.2	769	10	US-09-982-107-10	Sequence 10, Appl
25	17	2.2	771	12	US-09-969-748C-13	Sequence 13, Appl
26	16	2.1	771	10	US-09-818-247-4	Sequence 4, Appl
27	16	2.1	771	10	US-09-982-107-8	Sequence 8, Appl
28	13	1.7	733	10	US-09-818-247-5	Sequence 5, Appl
29	13	1.7	773	10	US-09-818-247-6	Sequence 6, Appl
30	13	1.7	773	10	US-09-982-107-2	Sequence 2, Appl
31	11	1.4	258	12	US-09-969-748C-111	Sequence 11, Appl
32	10	1.3	10	12	US-09-969-748C-74	Sequence 74, Appl
33	9	1.2	9	10	US-09-818-247-12	Sequence 12, Appl
34	9	1.2	624	11	US-09-491-322-22	Sequence 22, Appl
35	8	1.0	106	9	US-09-864-761-43512	Sequence 43512, A
36	8	1.0	140	10	US-09-970-088-11	Sequence 11, Appl
37	8	1.0	140	10	US-09-970-088-13	Sequence 13, Appl
38	8	1.0	140	10	US-09-970-088-14	Sequence 14, Appl
39	8	1.0	777	15	US-10-105-801-3	Sequence 3, Appl
40	8	1.0	802	9	US-09-758-386-3	Sequence 3, Appl
41	8	1.0	802	12	US-10-354-358-44	Sequence 44, Appl
42	8	1.0	802	15	US-10-011-548-33	Sequence 33, Appl
43	8	1.0	1298	10	US-09-982-610-33	Sequence 33, Appl
44	8	1.0	1298	15	US-10-262-538-32	Sequence 32, Appl
45	8	1.0	1362	15	US-10-105-901-33	Sequence 33, Appl

#### ALIGNMENTS

RESULT 1  
US-09-818-247-1  
Sequence 1, Application US/09818247  
Patent No. US20020102657A1  
GENERAL INFORMATION:  
APPLICANT: Mostov, Keith E.  
APPLICANT: Chaplin, Steven J.  
APPLICANT: Richman-Eisenstat, Janice  
TITLE OF INVENTION: The Regents of the University of California  
TITLE OF INVENTION: Ligands Directed to the No. US20020102657A1-Stalk Region of p19R and Methods of Use  
FILE REFERENCE: 18062E-000910US  
CURRENT APPLICATION NUMBER: US/09/818, 247  
CURRENT FILING DATE: 2001-03-26  
PRIOR APPLICATION NUMBER: WO PCT/US01/09659  
PRIOR FILING DATE: 2001-03-26  
PRIOR APPLICATION NUMBER: US 60/192,197  
PRIOR FILING DATE: 2000-03-27  
PRIOR APPLICATION NUMBER: US 60/192,198  
PRIOR FILING DATE: 2000-03-27  
NUMBER OF SEQ ID NOS: 26  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 1  
LENGTH: 764  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:  
OTHER INFORMATION: human polynuclear immunoglobulin receptor (p19R)  
US-09-818-247-1  
Query Match 100.0%; Score 764; DB 10; Length 764;  
Best Local Similarity 100.0%; Pred. NO. 0;  
Matches 764; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Oy 1 MLFLVTLCLAVFAIPSTKSPIFGPEEVNSVSGNSVSTICYPPTSVNRHRTKYCRGA 60  
DB 1 MLFLVTLCLAVFAIPSTKSPIFGPEEVNSVSGNSVSTICYPPTSVNRHRTKYCRGA 60

QY 61 RGCITLISSEGVSSKYAGRANLTNPENGTFVNNIAQLSODDSGRYKCGGLNSRGLS 120  
DB 61 RGCITLISSEGVSSKYAGRANLTNPENGTFVNNIAQLSODDSGRYKCGGLNSRGLS 120  
QY 121 FDVLSLEVSOGPGLNLTNTKYTVDLGRTVTINCPFKTENAKRKSLYKQIGLYPVLYDSS 180  
DB 121 FDVLSLEVSOGPGLNLTNTKYTVDLGRTVTINCPFKTENAKRKSLYKQIGLYPVLYDSS 180  
QY 181 GYVNPVYTGRIRLDIOGTGQLFSVYINQLRLSDAQOYLCOAGDSDNSNKKNDLOYLKP 240  
DB 181 GYVNPVYTGRIRLDIOGTGQLFSVYINQLRLSDAQOYLCOAGDSDNSNKKNDLOYLKP 240  
QY 241 EPELVYEDLRGSYTFHCALGPEVANAFLCROSSGSCNCDVYVNTLGKRAPAFEGRTLLN 300  
DB 241 EPELVYEDLRGSYTFHCALGPEVANAFLCROSSGSCNCDVYVNTLGKRAPAFEGRTLLN 300  
QY 301 PODKDSFSVITGLRKEDAGRYLCGAHSDGOLOEGSPIQAMOLFVNEESTIPRSPVYK 360  
DB 301 PODKDSFSVITGLRKEDAGRYLCGAHSDGOLOEGSPIQAMOLFVNEESTIPRSPVYK 360  
QY 361 GVAOSSVAVLCPYNNRRESKSIKYWCLMEGAQNGRCPLVDSEGMVKAQYEGRLSLEEPC 420  
DB 361 GVAOSSVAVLCPYNNRRESKSIKYWCLMEGAQNGRCPLVDSEGMVKAQYEGRLSLEEPC 420  
QY 421 NGFTYILNOLTSRDAGFYWCLTNGDTLMTVEIKIIEGEPNLKVPNGNTAVLGETLKY 480  
DB 421 NGFTYILNOLTSRDAGFYWCLTNGDTLMTVEIKIIEGEPNLKVPNGNTAVLGETLKY 480  
QY 481 PCHEPKCFSSYEYKWKMNNTGCOALPSODEGSKAFVNCDENSRLVLTNLVTRADG 540  
DB 481 PCHEPKCFSSYEYKWKMNNTGCOALPSODEGSKAFVNCDENSRLVLTNLVTRADG 540  
QY 541 WYWCYKQOGHFYGETAAVYVAEERKAAGSRDVS LAKADAPEKYLDSGFREIENKAIO 600  
DB 541 WYWCYKQOGHFYGETAAVYVAEERKAAGSRDVS LAKADAPEKYLDSGFREIENKAIO 600  
QY 601 DPLFLAEKAVADTPOADSRASVSGSSEBOGSSRALVSTLVLGLVLAAGAAYGV 660  
DB 601 DPLFLAEKAVADTPOADSRASVSGSSEBOGSSRALVSTLVLGLVLAAGAAYGV 660  
QY 661 ARAHRKNVDVRSIRSYRTDISMSDFENSREFGANDMGASSITTOETSLGKEEFVATTE 720  
DB 661 ARAHRKNVDVRSIRSYRTDISMSDFENSREFGANDMGASSITTOETSLGKEEFVATTE 720  
QY 721 STTEKEPKKAKRSSKEEAMAYKDFLLOSTVAAEAQDGPQEA 764  
DB 721 STTEKEPKKAKRSSKEEAMAYKDFLLOSTVAAEAQDGPQEA 764

RESULT 2

US-09-981-353-59  
; Sequence 59, Application US/09981353  
; Patent No. US20020160382A1  
; GENERAL INFORMATION:  
; APPLICANT: Lasek, Amy W.  
; TITLE OF INVENTION: GENES EXPRESSED IN COLON CANCER  
; FILE REFERENCE: PA-0038 US  
; CURRENT APPLICATION NUMBER: US/09/981.353  
; CURRENT FILING DATE: 2001-10-11  
; NUMBER OF SEQ ID NOS: 194  
; SOFTWARE: PERL Program  
; SEQ ID NO 59  
; LENGTH: 764  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc\_feature  
; OTHER INFORMATION: Incyte ID No. US20020160382A1 2680109CD1  
US-09-981-353-59

Query Match 100.0%; Score 764; DB 10; Length 764;  
Best Local Similarity 100.0%; Pred. No. 0;

Matches 764; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MLLFVLTCLLAVPAISTSPIFGPEEVNSVEGNSVITCYXPTSVNRRHTRKYWCRGA 60  
DB 1 MLLFVLTCLLAVPAISTSPIFGPEEVNSVEGNSVITCYXPTSVNRRHTRKYWCRGA 60  
QY 61 RGCITLISSEGVSSKYAGRANLTNPENGTFVNNIAQLSODDSGRYKCGGLNSRGLS 120  
DB 61 RGCITLISSEGVSSKYAGRANLTNPENGTFVNNIAQLSODDSGRYKCGGLNSRGLS 120  
QY 121 FDVLSLEVSOGPGLNLTNTKYTVDLGRTVTINCPFKTENAKRKSLYKQIGLYPVLYDSS 180  
DB 121 FDVLSLEVSOGPGLNLTNTKYTVDLGRTVTINCPFKTENAKRKSLYKQIGLYPVLYDSS 180  
QY 181 GYVNPVYTGRIRLDIOGTGQLFSVYINQLRLSDAQOYLCOAGDSDNSNKKNDLOYLKP 240  
DB 181 GYVNPVYTGRIRLDIOGTGQLFSVYINQLRLSDAQOYLCOAGDSDNSNKKNDLOYLKP 240  
QY 241 EPELVYEDLRGSYTFHCALGPEVANAFLCROSSGSCNCDVYVNTLGKRAPAFEGRTLLN 300  
DB 241 EPELVYEDLRGSYTFHCALGPEVANAFLCROSSGSCNCDVYVNTLGKRAPAFEGRTLLN 300  
QY 301 PODKDSFSVITGLRKEDAGRYLCGAHSDGOLOEGSPIQAMOLFVNEESTIPRSPVYK 360  
DB 301 PODKDSFSVITGLRKEDAGRYLCGAHSDGOLOEGSPIQAMOLFVNEESTIPRSPVYK 360  
QY 361 GVAOSSVAVLCPYNNRRESKSIKYWCLMEGAQNGRCPLVDSEGMVKAQYEGRLSLEEPC 420  
DB 361 GVAOSSVAVLCPYNNRRESKSIKYWCLMEGAQNGRCPLVDSEGMVKAQYEGRLSLEEPC 420  
QY 421 NGFTYILNOLTSRDAGFYWCLTNGDTLMTVEIKIIEGEPNLKVPNGNTAVLGETLKY 480  
DB 421 NGFTYILNOLTSRDAGFYWCLTNGDTLMTVEIKIIEGEPNLKVPNGNTAVLGETLKY 480  
QY 481 PCHEPKCFSSYEYKWKMNNTGCOALPSODEGSKAFVNCDENSRLVLTNLVTRADG 540  
DB 481 PCHEPKCFSSYEYKWKMNNTGCOALPSODEGSKAFVNCDENSRLVLTNLVTRADG 540  
QY 541 WYWCYKQOGHFYGETAAVYVAEERKAAGSRDVS LAKADAPEKYLDSGFREIENKAIO 600  
DB 541 WYWCYKQOGHFYGETAAVYVAEERKAAGSRDVS LAKADAPEKYLDSGFREIENKAIO 600  
QY 601 DPLFLAEKAVADTPOADSRASVSGSSEBOGSSRALVSTLVLGLVLAAGAAYGV 660  
DB 601 DPLFLAEKAVADTPOADSRASVSGSSEBOGSSRALVSTLVLGLVLAAGAAYGV 660  
QY 661 ARAHRKNVDVRSIRSYRTDISMSDFENSREFGANDMGASSITTOETSLGKEEFVATTE 720  
DB 661 ARAHRKNVDVRSIRSYRTDISMSDFENSREFGANDMGASSITTOETSLGKEEFVATTE 720  
QY 721 STTEKEPKKAKRSSKEEAMAYKDFLLOSTVAAEAQDGPQEA 764  
DB 721 STTEKEPKKAKRSSKEEAMAYKDFLLOSTVAAEAQDGPQEA 764

RESULT 3

US-09-989-919-124  
; Sequence 124, Application US/09989919  
; Patent No. US20020164344A1  
; GENERAL INFORMATION:  
; APPLICANT: Macina, Roberto  
; APPLICANT: Recipon, Herve  
; APPLICANT: Pluta, Jason  
; APPLICANT: Ghosh, Malavika  
; APPLICANT: Sun, Yongming  
; APPLICANT: Liu, Chenghua  
; TITLE OF INVENTION: Compositions and Methods Relating to Colon Specific Genes at  
; FILE REFERENCE: DEX-0289  
; CURRENT APPLICATION NUMBER: US/09/989.919  
; CURRENT FILING DATE: 2001-11-21  
; PRIOR APPLICATION NUMBER: 60/252,505  
; NUMBER OF SEQ ID NOS: 124

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 124  
; LENGTH: 764  
; TYPE: PRT  
; ORGANISM: Homo sapien  
US-09-989-919-124Query Match 100.0%; Score 764; DB 10; Length 764;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 764; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLFLVLTCLLAFPAISTKSPFPGPEEVNSVEGNSVSTICYPPTSVNHTTKYWCROGA 60  
DB 1 MLFLVLTCLLAFPAISTKSPFPGPEEVNSVEGNSVSTICYPPTSVNHTTKYWCROGA 60  
QY 61 RGCCTLLISSEGVSSKYAGRANLTFNPFNGTFVYVNIADLSODDSGRYKCGGINSRGLS 120  
DB 61 RGCCTLLISSEGVSSKYAGRANLTFNPFNGTFVYVNIADLSODDSGRYKCGGINSRGLS 120  
QY 121 FDVSLSEVSGPGLNDTKVYTVDLGRVTYINCPFTENAKRSLYKQIGLYPVLVIDSS 180  
DB 121 FDVSLSEVSGPGLNDTKVYTVDLGRVTYINCPFTENAKRSLYKQIGLYPVLVIDSS 180  
QY 181 GYVNPNTGRIKRLDIOGTGQLLFVYINQLRLSDAGQYLCQAGDSDNSNKKKADLOVLP 240  
DB 181 GYVNPNTGRIKRLDIOGTGQLLFVYINQLRLSDAGQYLCQAGDSDNSNKKKADLOVLP 240  
QY 241 EPELYEDLRGSVTFHCALGPEVANVAKFLCRSGSGENCDDVYVNTLGKRAPFEGRIILN 300  
DB 241 EPELYEDLRGSVTFHCALGPEVANVAKFLCRSGSGENCDDVYVNTLGKRAPFEGRIILN 300  
QY 301 PODKDGFSVYITGLRKEADAGRYLCGAHSDGQLOEGSPLOAMOLFVNEESTIPRSPTYVK 360  
DB 301 PODKDGFSVYITGLRKEADAGRYLCGAHSDGQLOEGSPLOAMOLFVNEESTIPRSPTYVK 360  
QY 361 GVAGSSVAVLCPYNNKESKSTIKYKLMGCAQNGRCPLVDSEGMWKAQYEGRLSLEEPG 420  
DB 361 GVAGSSVAVLCPYNNKESKSTIKYKLMGCAQNGRCPLVDSEGMWKAQYEGRLSLEEPG 420  
QY 421 NGTFVILNQLTSRDAGFYWCLTNGDTLMTTVEIKIIEGEPNLKYPGNVTVLGEFTLKV 480  
DB 421 NGTFVILNQLTSRDAGFYWCLTNGDTLMTTVEIKIIEGEPNLKYPGNVTVLGEFTLKV 480  
QY 481 PCHFPCKFSSEYKCYCKNNNTGCOALPDSODEGSPKAFVNCDENSRLVSTLTLNLTAD 540  
DB 481 PCHFPCKFSSEYKCYCKNNNTGCOALPDSODEGSPKAFVNCDENSRLVSTLTLNLTAD 540  
QY 541 WYMGCVKGHFGEETAAVYVAVEERKAAGSRDVSIAKADAPDEKVLDSGFREIENKAIQ 600  
DB 541 WYMGCVKGHFGEETAAVYVAVEERKAAGSRDVSIAKADAPDEKVLDSGFREIENKAIQ 600  
QY 601 DPLFAEKAADVADTRDQADGSRASVDSGSSSEEGSSRALVSTLPLGLVLAAGVAVGV 660  
DB 601 DPLFAEKAADVADTRDQADGSRASVDSGSSSEEGSSRALVSTLPLGLVLAAGVAVGV 660  
QY 661 ARAHRKRVNDRVYSIRSYRTDISMSDFENSREFGANDNKGASSITQETSLGKEEFAVTE 720  
DB 661 ARAHRKRVNDRVYSIRSYRTDISMSDFENSREFGANDNKGASSITQETSLGKEEFAVTE 720  
QY 721 STTETKEPKKAKRSSKEEAMAYKDFLLQSSSTVAEAODGPOEA 764  
DB 721 STTETKEPKKAKRSSKEEAMAYKDFLLQSSSTVAEAODGPOEA 764

RESULT 4  
US-09-982-107-4; Sequence 4, Application US/09982107  
; Patent No. US2002015958A1  
; GENERAL INFORMATION:  
; APPLICANT: HIATT, ANDREW C.  
; APPLICANT: HEIN, MICHAEL B.  
; TITLE OF INVENTION: METHODS FOR PRODUCING IMMUNOGLOBULINS CONTAINING  
; TITLE OF INVENTION: PROTECTION PROTEINS IN PLANTS AND THEIR USE

; FILE REFERENCE: EP13002E

; CURRENT APPLICATION NUMBER: US/09/982.107  
; CURRENT FILING DATE: 2001-10-16  
; NUMBER OF SEQ ID NOS: 19  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 4  
; LENGTH: 746  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-982-107-4Query Match 97.6%; Score 746; DB 10; Length 746;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 746; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 19 KSPIGPEEVNSVEGNSVSTICYPPTSVNHTTKYWCROGARGCITLLISSEGVSSKY 78  
DB 1 KSPIGPEEVNSVEGNSVSTICYPPTSVNHTTKYWCROGARGCITLLISSEGVSSKY 78  
QY 79 AGRANLTFNPFNGTFVYVNIADLSODDSGRYKCGGINSRGLSFDVSLSEVSGPGLNDTK 138  
DB 79 AGRANLTFNPFNGTFVYVNIADLSODDSGRYKCGGINSRGLSFDVSLSEVSGPGLNDTK 138  
QY 139 VYTVDLGRVTYINCPFTENAKRSLYKQIGLYPVLVIDSSGYVNPNTGRIKRLDIOGT 180  
DB 139 VYTVDLGRVTYINCPFTENAKRSLYKQIGLYPVLVIDSSGYVNPNTGRIKRLDIOGT 180  
QY 199 GOLLFSVYINQLRLSDAGQYLCQAGDSDNSNKKKADLOVLPPELYEDLRGSVTFHCA 258  
DB 199 GOLLFSVYINQLRLSDAGQYLCQAGDSDNSNKKKADLOVLPPELYEDLRGSVTFHCA 258  
QY 240 COLLEFSVYINQLRLSDAGQYLCQAGDSDNSNKKKADLOVLPPELYEDLRGSVTFHCA 240  
DB 240 COLLEFSVYINQLRLSDAGQYLCQAGDSDNSNKKKADLOVLPPELYEDLRGSVTFHCA 240  
QY 259 LGPEVANVAKFLCRSGSGENCDDVYVNTLGKRAPFEGRIILNPODKGFSVYITGLRKE 318  
DB 259 LGPEVANVAKFLCRSGSGENCDDVYVNTLGKRAPFEGRIILNPODKGFSVYITGLRKE 318  
QY 319 DAGRLTCAHSDGQLOEGSPLOAMOLFVNEESTIPRSPTYVKGVAGSSVAVLCPYNNKES 378  
DB 319 DAGRLTCAHSDGQLOEGSPLOAMOLFVNEESTIPRSPTYVKGVAGSSVAVLCPYNNKES 378  
QY 361 DAGRYLCGAHSDGQLOEGSPLOAMOLFVNEESTIPRSPTYVKGVAGSSVAVLCPYNNKES 360  
DB 361 DAGRYLCGAHSDGQLOEGSPLOAMOLFVNEESTIPRSPTYVKGVAGSSVAVLCPYNNKES 360  
QY 379 KSIRKWMCEGAQNGRCPLVDSEGMWKAQYEGRLSLEEPGNGTFVILNQLTSRDAGF 438  
DB 379 KSIRKWMCEGAQNGRCPLVDSEGMWKAQYEGRLSLEEPGNGTFVILNQLTSRDAGF 438  
QY 439 KSIRKWMCEGAQNGRCPLVDSEGMWKAQYEGRLSLEEPGNGTFVILNQLTSRDAGF 438  
DB 439 KSIRKWMCEGAQNGRCPLVDSEGMWKAQYEGRLSLEEPGNGTFVILNQLTSRDAGF 438  
QY 481 WYMGCVKGHFGEETAAVYVAVEERKAAGSRDVSIAKADAPDEKVLDSGFREIENKAIQ 540  
DB 481 WYMGCVKGHFGEETAAVYVAVEERKAAGSRDVSIAKADAPDEKVLDSGFREIENKAIQ 540  
QY 541 WYMGCVKGHFGEETAAVYVAVEERKAAGSRDVSIAKADAPDEKVLDSGFREIENKAIQ 600  
DB 541 WYMGCVKGHFGEETAAVYVAVEERKAAGSRDVSIAKADAPDEKVLDSGFREIENKAIQ 600  
QY 601 DPLFAEKAADVADTRDQADGSRASVDSGSSSEEGSSRALVSTLPLGLVLAAGVAVGV 660  
DB 601 DPLFAEKAADVADTRDQADGSRASVDSGSSSEEGSSRALVSTLPLGLVLAAGVAVGV 660  
QY 661 ARAHRKRVNDRVYSIRSYRTDISMSDFENSREFGANDNKGASSITQETSLGKEEFAVTE 720  
DB 661 ARAHRKRVNDRVYSIRSYRTDISMSDFENSREFGANDNKGASSITQETSLGKEEFAVTE 720  
QY 721 STTETKEPKKAKRSSKEEAMAYKDFLLQSSSTVAEAODGPOEA 764  
DB 721 STTETKEPKKAKRSSKEEAMAYKDFLLQSSSTVAEAODGPOEA 764

RESULT 5  
US-09-950-294-4; Sequence 4, Application US/09950294  
; Patent No. US20020127645A1  
; GENERAL INFORMATION:

APPLICANT: Morrison, Sherie L.  
Chinlatchavuvu, Kote R.  
TITLE OF INVENTION: SECRETORY IMMUNOGLOBULIN PRODUCED  
BY SINGLE CELLS AND METHODS FOR MAKING AND USING  
SAME  
NUMBER OF SEQUENCES: 4  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Merchant, Gould, Smith, Edell, Welter & Schmidt  
STREET: 11150 Santa Monica Boulevard, Suite 400  
City: Los Angeles  
STATE: CA  
COUNTRY: USA  
ZIP: 90025  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FASTSEQ for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/950,294  
FILING DATE: 10-Sep-2001  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 09/095,385  
FILING DATE: <Unknown>  
ATTORNEY/AGENT INFORMATION:  
NAME: Canada, Karen S  
REGISTRATION NUMBER: 39,927  
REFERENCE/DOCKET NUMBER: 30435,45USU1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 310 445-1140  
TELEFAX: 310 445-9031  
TELEX: <Unknown>  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 608 amino acids  
TYPE: amino acid  
STRANDEDNESS: unknown  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 4:  
US-09-950-294-4  
Query Match 79.5%; Score 607; DB 10; Length 608;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 607; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MLPLVLTCLAVPALTSTKSPITGPEEVNSVEGNSVITCYVPTSVNRRTRKRYMCRGA 60  
DB 1 MLPLVLTCLAVPALTSTKSPITGPEEVNSVEGNSVITCYVPTSVNRRTRKRYMCRGA 60  
QY 61 RGCITLISSEGVSSKYAGRANLTNPENGTFVNNIAQLSDDSGRKGLGINSRGLS 120  
DB 61 RGCITLISSEGVSSKYAGRANLTNPENGTFVNNIAQLSDDSGRKGLGINSRGLS 120  
QY 121 FDVLSLEVSGPGLNDTKYTVLGRVTINCEPKTENAKRSLKQIOLPYLVYDSS 180  
DB 121 FDVLSLEVSGPGLNDTKYTVLGRVTINCEPKTENAKRSLKQIOLPYLVYDSS 180  
QY 181 GYVNPNTGRIRLDIOGTGLLFSVYINQLRLSDAGGYLCOAGDDSNKKNADLOVLRP 240  
DB 181 GYVNPNTGRIRLDIOGTGLLFSVYINQLRLSDAGGYLCOAGDDSNKKNADLOVLRP 240  
QY 241 EPELVYEDLRGVTFHGALGPVYANNAKFLCROSSGENDCVVNTLGRAPAREGRILLN 300  
DB 241 EPELVYEDLRGVTFHGALGPVYANNAKFLCROSSGENDCVVNTLGRAPAREGRILLN 300  
QY 301 POKKGSFSVYITGLRKEDAGRYLCAHSDGQLOEGSPIAQLFVNEESTIRSPVYVK 360  
DB 301 POKKGSFSVYITGLRKEDAGRYLCAHSDGQLOEGSPIAQLFVNEESTIRSPVYVK 360  
QY 361 GVAGSSVAVLCPTNRRKSKIKYWCLEGAQNGRCPLVDSEGKVKQYGRISLLEPG 420

DB 361 GVAGSSVAVLCPTNRRKSKIKYWCLEGAQNGRCPLVDSEGKVKQYGRISLLEPG 420  
QY 421 NGTFYILNOLTSRDAFFYWCILNGLDGLMTVEIKIIEGPNLKPAGNVTAVGETLKV 480  
DB 421 NGTFYILNOLTSRDAFFYWCILNGLDGLMTVEIKIIEGPNLKPAGNVTAVGETLKV 480  
QY 481 PCHEPCKFSSEYKWKWNTGCOALPSODEGSKAFVNCDENSRLVSLTNLYTRADEG 540  
DB 481 PCHEPCKFSSEYKWKWNTGCOALPSODEGSKAFVNCDENSRLVSLTNLYTRADEG 540  
QY 541 WYMGVKGQHFYGETAAVYAVEERKAGSRDVSLLAKADAPDEKVLDSGFEIENKAIO 600  
DB 541 WYMGVKGQHFYGETAAVYAVEERKAGSRDVSLLAKADAPDEKVLDSGFEIENKAIO 600  
QY 601 DPLRFAE 607  
DB 601 DPLRFAE 607  
RESULT 6  
US-09-969-748C-108  
Sequence 108, Application US/09969748C  
Publication No. US20030161809A1  
GENERAL INFORMATION:  
APPLICANT: ARIZENE PHARMACEUTICALS, INC.  
APPLICANT: HOUSTON, Lou, L.  
APPLICANT: SHERIDAN, Philip, J.  
APPLICANT: HAWLEY, Stephen  
APPLICANT: GLYNN, Jacqueline, M.  
APPLICANT: CHAPIN, Steven  
APPLICANT: BASU, Amresh  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE TRANSPORT OF BIOLOGICALLY  
FILE REFERENCE: 057220-0303  
CURRENT APPLICATION NUMBER: US/09/969,748C  
CURRENT FILING DATE: 2002-12-10  
PRIOR APPLICATION NUMBER: US 60/267,601  
PRIOR FILING DATE: 2001-02-09  
PRIOR APPLICATION NUMBER: US 60/248,819  
PRIOR FILING DATE: 2000-11-14  
PRIOR APPLICATION NUMBER: US 60/248,478  
PRIOR FILING DATE: 2000-11-13  
PRIOR APPLICATION NUMBER: US 60/237,929  
PRIOR FILING DATE: 2000-10-02  
NUMBER OF SEQ ID NOS: 115  
SOFTWARE: PatentIn version 3.0  
SEQ ID NO 108  
LENGTH: 243  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-969-748C-108  
Query Match 31.8%; Score 243; DB 12; Length 243;  
Best Local Similarity 100.0%; Pred. No. 1,3e-232;  
Matches 243; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 492 EKYCKWNTGCOALPSODEGSKAFVNCDENSRLVSLTNLYTRADEGKWKVQGHF 551  
DB 1 EKYCKWNTGCOALPSODEGSKAFVNCDENSRLVSLTNLYTRADEGKWKVQGHF 60  
QY 552 YGETAAVYAVEERKAGSRDVSLLAKADAPDEKVLDSGFEIENKAIOPLFAEKKAV 611  
DB 552 YGETAAVYAVEERKAGSRDVSLLAKADAPDEKVLDSGFEIENKAIOPLFAEKKAV 611  
QY 612 ADTRDADGSRASVDSGSEEOGSSRALVSTLPLGLVLAVGAVAVGARRHRRKNVDR 671  
DB 121 ADTRDADGSRASVDSGSEEOGSSRALVSTLPLGLVLAVGAVAVGARRHRRKNVDR 180  
QY 672 VSTRSYRTDIMSDFENSRFGANDMKAGSITOEYSLGKEEFVATTSSTETTKPKKA 731  
DB 181 VSTRSYRTDIMSDFENSRFGANDMKAGSITOEYSLGKEEFVATTSSTETTKPKKA 240  
QY 732 KRS 734

Db 241 KRS 243

# RESULT 7

US-10-047-542-51

Sequence 51, Application US/10047542  
 Publication No. US20020168367A1  
 GENERAL INFORMATION:  
 APPLICANT: LARRICK, JAMES W.  
 APPLICANT: WYCOFF, KEITH L.  
 TITLE OF INVENTION: NOVEL IMMUNOADHESINS FOR TREATING AND PREVENTING VIRAL  
 FILE REFERENCE: 030905.0004.C1P1  
 CURRENT APPLICATION NUMBER: US/10/047,542  
 PRIOR FILING DATE: 2001-10-26  
 PRIOR APPLICATION NUMBER: PCT/US01/13932  
 PRIOR FILING DATE: 2001-04-28  
 PRIOR APPLICATION NUMBER: 60/200,298  
 NUMBER OF SEQ ID NOS: 101  
 SOFTWARE: PatentIn ver. 2.1  
 SEQ ID NO 51  
 LENGTH: 602  
 TYPE: PRT  
 ORGANISM: Homo sapiens  
 US-10-047-542-51

Query Match 26.8%; Score 205; DB 14; Length 602;

Best Local Similarity 100.0%; Pred. No. 1.6e-194; Mismatches 0; Indels 0; Gaps 0;

Matches 205; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 51 TRKYCRGARGCGCITLISSEGYSSKYAGRANLTNPENCTFVNIAQLSQDDSGRYKC 110  
 DB 50 TRKYCRGARGCGCITLISSEGYSSKYAGRANLTNPENCTFVNIAQLSQDDSGRYKC 109  
 QY 111 GLGINSRGLSFDSVLEYSQCGELLNDTKVTVVDIGRTVINCPTENAKRKSLYQIG 170  
 DB 110 GLGINSRGLSFDSVLEYSQCGELLNDTKVTVVDIGRTVINCPTENAKRKSLYQIG 169  
 QY 171 LYPVLVIDSSGYVPNTGTRRLDIOGTGOLFESVVIINQLRLSDAGOTLCOAGDSSNSK 230  
 DB 170 LYPVLVIDSSGYVPNTGTRRLDIOGTGOLFESVVIINQLRLSDAGOTLCOAGDSSNSK 229  
 QY 231 KNADLOVLRKPEPELVYEDLRGSYTF 255  
 DB 230 KNADLOVLRKPEPELVYEDLRGSYTF 254

## RESULT 8

US-09-818-247-20

Sequence 20, Application US/09818247  
 Patent No. US20020102657A1  
 GENERAL INFORMATION:  
 APPLICANT: Mostov, Keith E.  
 APPLICANT: Chaplin, Steven J.  
 APPLICANT: Richman-Eisenstat, Janice  
 TITLE OF INVENTION: The Regents of the University of California  
 TITLE OF INVENTION: Ligands directed to the No. US20020102657A1-Sstalk Region of p19R and Methods of Use Ther  
 FILE REFERENCE: 18062E-000910US  
 CURRENT APPLICATION NUMBER: US/09/818,247  
 PRIOR FILING DATE: 2001-03-26  
 PRIOR APPLICATION NUMBER: WO PCT/US01/09699  
 PRIOR FILING DATE: 2001-03-26  
 PRIOR APPLICATION NUMBER: US 60/192,197  
 PRIOR FILING DATE: 2000-03-27  
 PRIOR APPLICATION NUMBER: US 60/192,198  
 PRIOR FILING DATE: 2000-03-27  
 NUMBER OF SEQ ID NOS: 26  
 SOFTWARE: PatentIn ver. 2.1  
 SEQ ID NO 20  
 LENGTH: 95

TYPE: PRT  
 ORGANISM: Artificial Sequence  
 FEATURE:  
 OTHER INFORMATION: Description of Artificial Sequence: portion of  
 OTHER INFORMATION: human p19R  
 US-09-818-247-20

## Query Match

Best Local Similarity 12.4%; Score 95; DB 10; Length 95;

Matches 95; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 544 CGVKGHFYGETAAYVAVEERKAGSRDVSIAKADAAPDEKVIDSGFREIENKAIDPPR 603  
 DB 1 CGVKGHFYGETAAYVAVEERKAGSRDVSIAKADAAPDEKVIDSGFREIENKAIDPPR 60  
 QY 604 LFAEKAVIDTRDQADGSRASVDSGSSSEEGGSSR 638  
 DB 61 LFAEKAVIDTRDQADGSRASVDSGSSSEEGGSSR 95

## RESULT 9

US-09-969-748C-104

Sequence 104, Application US/09969748C  
 Publication No. US20030161809A1  
 GENERAL INFORMATION:  
 APPLICANT: ARIZKE, PHARMACEUTICALS, INC.  
 APPLICANT: HOUSTON, Lou, L.  
 APPLICANT: SHERIDAN, Phillip, J.  
 APPLICANT: HAWLEY, Stephen  
 APPLICANT: GLYNN, Jacqueline, M.  
 APPLICANT: CHAPIN, Steven  
 APPLICANT: BASU, Amarendra  
 TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE TRANSPORT OF BIOLOGICALLY  
 FILE REFERENCE: 057220-0303  
 CURRENT APPLICATION NUMBER: US/09/969,748C  
 PRIOR FILING DATE: 2002-12-10  
 PRIOR APPLICATION NUMBER: US 60/267,601  
 PRIOR FILING DATE: 2001-02-09  
 PRIOR APPLICATION NUMBER: US 60/248,819  
 PRIOR FILING DATE: 2000-11-14  
 PRIOR APPLICATION NUMBER: US 60/248,478  
 PRIOR FILING DATE: 2000-11-13  
 PRIOR APPLICATION NUMBER: US 60/237,929  
 PRIOR FILING DATE: 2000-10-02  
 NUMBER OF SEQ ID NOS: 115  
 SOFTWARE: PatentIn version 3.0  
 SEQ ID NO 104  
 LENGTH: 102  
 TYPE: PRT  
 ORGANISM: Homo sapiens  
 US-09-969-748C-104

Query Match 12.3%; Score 94; DB 12; Length 102;

Best Local Similarity 100.0%; Pred. No. 4.8e-85; Mismatches 0; Indels 0; Gaps 0;

Matches 94; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 545 GVKGHHFYGETAAYVAVEERKAGSRDVSIAKADAAPDEKVIDSGFREIENKAIDPPR 604  
 DB 3 GVKGHHFYGETAAYVAVEERKAGSRDVSIAKADAAPDEKVIDSGFREIENKAIDPPR 62  
 QY 605 LFAEKAVIDTRDQADGSRASVDSGSSSEEGGSSR 638  
 DB 63 LFAEKAVIDTRDQADGSRASVDSGSSSEEGGSSR 96

## RESULT 10

US-09-818-247-17

Sequence 17, Application US/09818247  
 Patent No. US20020102657A1  
 GENERAL INFORMATION:  
 APPLICANT: Mostov, Keith E.  
 APPLICANT: Chaplin, Steven J.

APPLICANT: Richman-Eisenstat, Janice  
TITLE OF INVENTION: The Regents of the University of California  
TITLE OF INVENTION: Ligands directed to the No. US20020102657A1-Secretory Component.  
FILE REFERENCE: No. US20020102657A1-Stalk Region of p19R and Methods of use Ther  
CURRENT APPLICATION NUMBER: 18062E-00091005  
CURRENT FILING DATE: 2001-03-26  
PRIOR APPLICATION NUMBER: US 60/192,197  
PRIOR FILING DATE: 2000-03-27  
PRIOR APPLICATION NUMBER: US 60/192,197  
PRIOR FILING DATE: 2000-03-27  
PRIOR FILING DATE: 2000-03-27  
NUMBER OF SEQ ID NOS: 26  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 17  
LENGTH: 90  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: portion of  
OTHER INFORMATION: human p19R encompassing part of domain 5 and  
OTHER INFORMATION: domain 6  
US-09-818-247-17

Query Match 11.8%; Score 90; DB 10; Length 90;  
Best Local Similarity 100.0%; Pred. No. 4e-81;  
Matches 90; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 549 GHEYGETAAVYAAVEERKKAAGSRDVLAKADAAPDEKVLDSGFEIENKAIDPRLFAEE 608  
DB 1 GHEYGETAAVYAAVEERKKAAGSRDVLAKADAAPDEKVLDSGFEIENKAIDPRLFAEE 60  
OY 609 KAAVATRDQADGSRASVDSGSSSEOGGSSR 638  
DB 61 KAAVATRDQADGSRASVDSGSSSEOGGSSR 90

RESULT 11  
US-09-925-299-1218  
Sequence 1218, Application US/09925299  
Patent No. US20020055627A1  
GENERAL INFORMATION:  
APPLICANT: Rosen et al.  
TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies  
FILE REFERENCE: PA102  
CURRENT APPLICATION NUMBER: US/09/925,299  
CURRENT FILING DATE: 2001-08-10  
PRIOR APPLICATION NUMBER: PCT/US00/05883  
PRIOR FILING DATE: 2000-03-08  
PRIOR APPLICATION NUMBER: 60/124,270  
PRIOR FILING DATE: 1999-03-12  
NUMBER OF SEQ ID NOS: 1556  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 1218  
LENGTH: 70  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: SITE  
LOCATION: (2)  
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
US-09-925-299-1218

Query Match 6.5%; Score 50; DB 9; Length 70;  
Best Local Similarity 100.0%; Pred. No. 1.6e-41;  
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 219 LCOAGDSSNNKKNADLOYLKPEPELYEDLRGSYTFHCALGPEVANVAK 268  
DB 3 LCOAGDSSNNKKNADLOYLKPEPELYEDLRGSYTFHCALGPEVANVAK 52

RESULT 12  
US-09-925-299-1218  
Sequence 1218, Application US/09925299  
Publication No. US20030040617A9  
GENERAL INFORMATION:  
APPLICANT: Rosen et al.  
TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies  
FILE REFERENCE: PA102  
CURRENT APPLICATION NUMBER: US/09/925,299  
CURRENT FILING DATE: 2001-08-10  
PRIOR APPLICATION NUMBER: PCT/US00/05883  
PRIOR FILING DATE: 2000-03-08  
PRIOR APPLICATION NUMBER: 60/124,270  
PRIOR FILING DATE: 1999-03-12  
NUMBER OF SEQ ID NOS: 1556  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 1218  
LENGTH: 70  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: SITE  
LOCATION: (2)  
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
US-09-925-299-1218

Query Match 6.5%; Score 50; DB 11; Length 70;  
Best Local Similarity 100.0%; Pred. No. 1.6e-41;  
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 219 LCOAGDSSNNKKNADLOYLKPEPELYEDLRGSYTFHCALGPEVANVAK 268  
DB 3 LCOAGDSSNNKKNADLOYLKPEPELYEDLRGSYTFHCALGPEVANVAK 52

RESULT 13  
US-09-969-748C-109  
Sequence 109, Application US/09969748C  
Publication No. US20030161809A1  
GENERAL INFORMATION:  
APPLICANT: ARIZONA PHARMACEUTICALS, INC.  
APPLICANT: HOUSTON, Lou, L.  
APPLICANT: SHERIDAN, Philip, J.  
APPLICANT: HAWLEY, Stephen  
APPLICANT: GLENN, Jacqueline, M.  
APPLICANT: CHAPIN, Steven  
APPLICANT: BASU, Amareesh  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE TRANSPORT OF BIOLOGICALLY  
TITLE OF INVENTION: AGENTS ACROSS CELLULAR BARRIERS  
FILE REFERENCE: 057220-0303  
CURRENT APPLICATION NUMBER: US/09/969,748C  
CURRENT FILING DATE: 2002-12-10  
PRIOR APPLICATION NUMBER: US 60/267,601  
PRIOR FILING DATE: 2001-02-09  
PRIOR APPLICATION NUMBER: US 60/248,819  
PRIOR FILING DATE: 2000-11-14  
PRIOR APPLICATION NUMBER: US 60/248,478  
PRIOR FILING DATE: 2000-11-13  
PRIOR APPLICATION NUMBER: US 60/237,929  
PRIOR FILING DATE: 2000-10-02  
NUMBER OF SEQ ID NOS: 115  
SOFTWARE: PatentIn version 3.0  
SEQ ID NO 109  
LENGTH: 244  
TYPE: PRT  
ORGANISM: Simian  
US-09-969-748C-109

Query Match 6.2%; Score 47; DB 12; Length 244;  
Best Local Similarity 100.0%; Pred. No. 4.9e-38;  
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 666 RKNVDRVSTIRYRTDISMSDFENSREFGANDKMGASSTQETSLGCK 712

Db 176 RKNDRVSRSTYRDISMSDFENSEFGANDMGASSTOETSLGK 222

RESULT 14  
US-09-818-247-19

Sequence 19, Application US/09818247  
Patent No. US20020102657A1

GENERAL INFORMATION:

APPLICANT: Mostov, Keith E.

APPLICANT: Chapin, Steven J.

APPLICANT: Richman-Eisenstat, Janice

TITLE OF INVENTION: The Regents of the University of California

TITLE OF INVENTION: Ligands Directed to the No. US20020102657A1-Stalk Region of p19R and Methods of Use Ther

FILE REFERENCE: 18062E-000910US

CURRENT FILING DATE: US/09/818,247

PRIOR APPLICATION NUMBER: US/09/818,247

PRIOR FILING DATE: 2001-03-26

PRIOR APPLICATION NUMBER: US 60/192,197

PRIOR FILING DATE: 2000-03-27

PRIOR APPLICATION NUMBER: US 60/192,198

PRIOR FILING DATE: 2000-03-27

NUMBER OF SEQ ID NOS: 26

SOFTWARE: PatentIn Ver. 2.1

SEQ ID NO 19

LENGTH: 31

TYPE: PRT

ORGANISM: Artificial Sequence

FEATURE: Description of Artificial Sequence: human p19R

OTHER INFORMATION: stalk

US-09-818-247-19

Query Match 4.1%: Score 31; DB 10; Length 31;

Best Local Similarity 100.0%: Pred. No. 5.6e-23;

Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 608 EKAVADTRDQADGSRASVDSGSSSEEGGSSR 638

Db 1 EKAVADTRDQADGSRASVDSGSSSEEGGSSR 31

RESULT 15  
US-09-969-748C-107

Sequence 107, Application US/09969748C

Publication No. US20030161809A1

GENERAL INFORMATION:

APPLICANT: ARIZEKE PHARMACEUTICALS, INC.

APPLICANT: HOUSTON, Lou, L.

APPLICANT: SHERIDAN, Philip, J.

APPLICANT: HAMLEY, Stephen

APPLICANT: GLYNN, Jacqueline, M.

APPLICANT: CHAPIN, Steven

TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE TRANSPORT OF BIOLOGICALLY ACTIVE

TITLE OF INVENTION: AGENTS ACROSS CELLULAR BARRIERS

FILE REFERENCE: 057220-0303

CURRENT APPLICATION NUMBER: US/09/969,748C

PRIOR APPLICATION NUMBER: US 60/267,601

PRIOR FILING DATE: 2001-02-09

PRIOR APPLICATION NUMBER: US 60/248,819

PRIOR FILING DATE: 2000-11-14

PRIOR APPLICATION NUMBER: US 60/248,478

PRIOR FILING DATE: 2000-11-13

PRIOR APPLICATION NUMBER: US 60/237,929

PRIOR FILING DATE: 2000-10-02

NUMBER OF SEQ ID NOS: 115

SOFTWARE: PatentIn version 3.0

SEQ ID NO 107

LENGTH: 94

TYPE: PRT  
ORGANISM: Artificial Sequence

FEATURE:

OTHER INFORMATION: consensus sequence

US-09-969-748C-107

Query Match 2.9%: Score 22; DB 12; Length 94;

Best Local Similarity 100.0%: Pred. No. 1.3e-13;

Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 545 GVKQGHFYGETAAYVAVEERK 566

Db 3 GVKQGHFYGETAAYVAVEERK 24

Search completed: September 9, 2003, 23:29:34

Job time : 67 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 9, 2003, 23:17:23 ; Search time 394 Seconds

(without alignments)  
1764.411 Million cell updates/sec

Title: US-09-818-247-1

Perfect score: 1 MLFVLITCLLAVFPAINSTKS.....DFLLQSTVAEAADGPOEA 764

Scoring table: OLIGO

Searched: 5728757 seqs, 909918778 residues

Word size: 0

Total number of hits satisfying chosen parameters: 5728757

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database:

Pending Patents-AA\_Main:\*

1: /cgn2\_6/ptodata/1/paa/PCTUS\_COMB.pep.\*  
2: /cgn2\_6/ptodata/1/paa/US06\_COMB.pep.\*  
3: /cgn2\_6/ptodata/1/paa/US07\_COMB.pep.\*  
4: /cgn2\_6/ptodata/1/paa/US08\_COMB.pep.\*  
5: /cgn2\_6/ptodata/1/paa/US081\_COMB.pep.\*  
6: /cgn2\_6/ptodata/1/paa/US082\_COMB.pep.\*  
7: /cgn2\_6/ptodata/1/paa/US083\_COMB.pep.\*  
8: /cgn2\_6/ptodata/1/paa/US084\_COMB.pep.\*  
9: /cgn2\_6/ptodata/1/paa/US085\_COMB.pep.\*  
10: /cgn2\_6/ptodata/1/paa/US086\_COMB.pep.\*  
11: /cgn2\_6/ptodata/1/paa/US087\_COMB.pep.\*  
12: /cgn2\_6/ptodata/1/paa/US088\_COMB.pep.\*  
13: /cgn2\_6/ptodata/1/paa/US089\_COMB.pep.\*  
14: /cgn2\_6/ptodata/1/paa/US090\_COMB.pep.\*  
15: /cgn2\_6/ptodata/1/paa/US091\_COMB.pep.\*  
16: /cgn2\_6/ptodata/1/paa/US092\_COMB.pep.\*  
17: /cgn2\_6/ptodata/1/paa/US093\_COMB.pep.\*  
18: /cgn2\_6/ptodata/1/paa/US094\_COMB.pep.\*  
19: /cgn2\_6/ptodata/1/paa/US095\_COMB.pep.\*  
20: /cgn2\_6/ptodata/1/paa/US096\_COMB.pep.\*  
21: /cgn2\_6/ptodata/1/paa/US097\_COMB.pep.\*  
22: /cgn2\_6/ptodata/1/paa/US097A\_COMB.pep.\*  
23: /cgn2\_6/ptodata/1/paa/US098\_COMB.pep.\*  
24: /cgn2\_6/ptodata/1/paa/US099\_COMB.pep.\*  
25: /cgn2\_6/ptodata/1/paa/US099B\_COMB.pep.\*  
26: /cgn2\_6/ptodata/1/paa/US100\_COMB.pep.\*  
27: /cgn2\_6/ptodata/1/paa/US101\_COMB.pep.\*  
28: /cgn2\_6/ptodata/1/paa/US102\_COMB.pep.\*  
29: /cgn2\_6/ptodata/1/paa/US103\_COMB.pep.\*  
30: /cgn2\_6/ptodata/1/paa/US104\_COMB.pep.\*  
31: /cgn2\_6/ptodata/1/paa/US106\_COMB.pep.\*  
32: /cgn2\_6/ptodata/1/paa/US106\_COMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	764	100.0	764	12 US-08-860-037-2	Sequence 2, Appl1

	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45		
	764	100.0	764	100.0	764	100.0	764	100.0	764	100.0	764	100.0	764	100.0	764	100.0	764	100.0	764	100.0	764	100.0	764	100.0	764	100.0	764	100.0	764	100.0	764	100.0	764	100.0	764	100.0	764	100.0	764	100.0	764	100.0	764	100.0	764	100.0
	17	17	22	23	24	24	25	25	20	21	25	32	32	25	16	20	20	25	26	22	22	28	28	25	28	32	32	25	25	22	22	22	22	24	25	24	13	70	24	32	32	22	22	28		
	US-09-316-412-30	US-09-791-537-92050	US-09-818-247-1	US-09-949-016-6254	US-09-949-039-65	US-09-981-353-59	US-09-989-919-124	US-09-989-919-124	US-09-982-107-4	US-09-982-107-4	US-09-982-107-4	US-09-982-107-4	US-09-982-107-4	US-09-982-107-4	US-09-982-107-4	US-09-982-107-4	US-09-982-107-4	US-09-982-107-4	US-09-982-107-4	US-09-982-107-4	US-09-982-107-4	US-09-982-107-4	US-09-982-107-4	US-09-982-107-4	US-09-982-107-4	US-09-982-107-4	US-09-982-107-4	US-09-982-107-4	US-09-982-107-4	US-09-982-107-4	US-09-982-107-4	US-09-982-107-4	US-09-982-107-4	US-09-982-107-4	US-09-982-107-4	US-09-982-107-4	US-09-982-107-4	US-09-982-107-4	US-09-982-107-4	US-09-982-107-4	US-09-982-107-4	US-09-982-107-4	US-09-982-107-4	US-09-982-107-4		
	Sequence 30, Appl1	Sequence 92050, A	Sequence 1, Appl1	Sequence 6254, Ap	Sequence 59, Appl	Sequence 124, Appl	Sequence 59, Appl	Sequence 59, Appl	Sequence 4, Appl1	Sequence 4, Appl1	Sequence 4, Appl1	Sequence 4, Appl1	Sequence 4, Appl1	Sequence 4, Appl1	Sequence 4, Appl1	Sequence 4, Appl1	Sequence 4, Appl1	Sequence 4, Appl1	Sequence 4, Appl1	Sequence 4, Appl1	Sequence 4, Appl1	Sequence 4, Appl1	Sequence 4, Appl1	Sequence 4, Appl1	Sequence 4, Appl1	Sequence 4, Appl1	Sequence 4, Appl1	Sequence 4, Appl1	Sequence 4, Appl1	Sequence 4, Appl1	Sequence 4, Appl1	Sequence 4, Appl1	Sequence 4, Appl1	Sequence 4, Appl1	Sequence 4, Appl1	Sequence 4, Appl1	Sequence 4, Appl1	Sequence 4, Appl1	Sequence 4, Appl1	Sequence 4, Appl1	Sequence 4, Appl1	Sequence 4, Appl1	Sequence 4, Appl1	Sequence 4, Appl1		

## ALIGNMENTS

RESULT 1  
US-08-860-037-2  
Sequence 2, Application US/08860037  
GENERAL INFORMATION:  
APPLICANT: Jan van Oostrum  
APPLICANT: Fredericus A.M. Asselbergs  
APPLICANT: Christian Leist  
APPLICANT: Markus Grueter  
TITLE OF INVENTION: Production of Recombinant Secretary  
TITLE OF INVENTION: Component  
NUMBER OF SEQUENCES: 7  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Michael Glynn, Novartis Corporation,  
STREET: 59 Route 10  
CITY: East Hanover  
STATE: New Jersey  
COUNTRY: U.S.  
ZIP: 07936  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/860,037  
FILING DATE: 11 June 1997  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/EP95/04797  
FILING DATE: 12 June 1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Furman, Diane E.  
REGISTRATION NUMBER: 31,104  
REFERENCE/DOCKET NUMBER: 4-20259/A/PCT  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 908-277-2311  
TELEFAX: 908-277-4009  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 764 amino acids  
TYPE: amino acid  
TOPOLOGY: 1linear  
MOLECULE TYPE: protein  
US-08-860-037-2

Query Match 100.0%; Score 764; DB 12; Length 764;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 764; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLFLVLTCLLAVPAISTKSPIFGPEEVNSVEGNSVITCYPTPTSVNHRTRKYKROGA 60  
DB 1 MLFLVLTCLLAVPAISTKSPIFGPEEVNSVEGNSVITCYPTPTSVNHRTRKYKROGA 60  
QY 61 RGGCITLISSEGVSSKYGARANTNFPENGTFVNNIAOLSODDSGRYKCGLINSRGLS 120  
DB 61 RGGCITLISSEGVSSKYGARANTNFPENGTFVNNIAOLSODDSGRYKCGLINSRGLS 120  
QY 121 FVLSLEVSGPGLNDTKYTYVDLGRVTINCPFKTENAKRKSLYKQIGLYPVLYIDSS 180  
DB 121 FVLSLEVSGPGLNDTKYTYVDLGRVTINCPFKTENAKRKSLYKQIGLYPVLYIDSS 180  
QY 181 GYVNPYTGRIRLDIOGTGQLLFVYINOLRLSDAQYICQAGDSDNSNKKNAADLYLKP 240  
DB 181 GYVNPYTGRIRLDIOGTGQLLFVYINOLRLSDAQYICQAGDSDNSNKKNAADLYLKP 240  
QY 241 EPELYEDLRGSVTFHCALGPEVANYAKFLCROSSGENCDVYVNTLGKRAPAFEGRIILN 300  
DB 241 EPELYEDLRGSVTFHCALGPEVANYAKFLCROSSGENCDVYVNTLGKRAPAFEGRIILN 300  
QY 301 PODKDSFSVITGLRKEDAGRYLCAHSDDGQLQEGSPIOAMQLFVNEESTIPRSPTVVK 360  
DB 301 PODKDSFSVITGLRKEDAGRYLCAHSDDGQLQEGSPIOAMQLFVNEESTIPRSPTVVK 360  
QY 361 GVAAGSSVAVLCYPNRRESKSIKWCIMBGAONRCPLVDSEGMVKAQYEGRLSLEEPC 420  
DB 361 GVAAGSSVAVLCYPNRRESKSIKWCIMBGAONRCPLVDSEGMVKAQYEGRLSLEEPC 420  
QY 421 NGFTVILNOLTSRDGFWCLNGDTLMRTTVEIKIIEGEPNLKYPGNTAVLGETLKY 480  
DB 421 NGFTVILNOLTSRDGFWCLNGDTLMRTTVEIKIIEGEPNLKYPGNTAVLGETLKY 480  
QY 481 PCHFPCKFSSYEKYKWKNNNTGCOALPSODEGSKAFVNCDENSRLVSLTLNLYTRADEG 540  
DB 481 PCHFPCKFSSYEKYKWKNNNTGCOALPSODEGSKAFVNCDENSRLVSLTLNLYTRADEG 540  
QY 541 WYMGCVKQGHFYGETAAYVAVEERKAAGSDVSLAKADAAPDEKVLDSGFREIENKAIO 600  
DB 541 WYMGCVKQGHFYGETAAYVAVEERKAAGSDVSLAKADAAPDEKVLDSGFREIENKAIO 600  
QY 601 DPLFLAEEKAVADTRDOAGSRASVDSGSSSEEGGSRALVSTLVLGLVLAAGAAYGV 660  
DB 601 DPLFLAEEKAVADTRDOAGSRASVDSGSSSEEGGSRALVSTLVLGLVLAAGAAYGV 660  
QY 661 ARAHRKKNDRVSIIRSYRTDISMSDFENSREFGANDMGASSITQETSLGKKEEVAATTE 720  
DB 661 ARAHRKKNDRVSIIRSYRTDISMSDFENSREFGANDMGASSITQETSLGKKEEVAATTE 720

QY 721 STTETKEPRKAKRSSKEEEMAYKDFLLQSSVTAEAODGPOEA 764  
DB 721 STTETKEPRKAKRSSKEEEMAYKDFLLQSSVTAEAODGPOEA 764

RESULT 2  
US-09-316-412-30  
Sequence 30, Application US/09316412  
GENERAL INFORMATION:  
APPLICANT: Sheppard, Paul O.  
TITLE OF INVENTION: IMMUNOMODULATOR POLYPEPTIDE ZSIG57  
FILE REFERENCE: 98-23  
CURRENT APPLICATION NUMBER: US/09/316,412  
CURRENT FILING DATE: 1999-05-20  
EARLIER FILING DATE: 1998-06-18  
NUMBER OF SEQ ID NOS: 30  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 30  
LENGTH: 764  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-316-412-30

Query Match 100.0%; Score 764; DB 17; Length 764;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 764; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLFLVLTCLLAVPAISTKSPIFGPEEVNSVEGNSVITCYPTPTSVNHRTRKYKROGA 60  
DB 1 MLFLVLTCLLAVPAISTKSPIFGPEEVNSVEGNSVITCYPTPTSVNHRTRKYKROGA 60  
QY 61 RGGCITLISSEGVSSKYGARANTNFPENGTFVNNIAOLSODDSGRYKCGLINSRGLS 120  
DB 61 RGGCITLISSEGVSSKYGARANTNFPENGTFVNNIAOLSODDSGRYKCGLINSRGLS 120  
QY 121 FVLSLEVSGPGLNDTKYTYVDLGRVTINCPFKTENAKRKSLYKQIGLYPVLYIDSS 180  
DB 121 FVLSLEVSGPGLNDTKYTYVDLGRVTINCPFKTENAKRKSLYKQIGLYPVLYIDSS 180  
QY 181 GYVNPYTGRIRLDIOGTGQLLFVYINOLRLSDAQYICQAGDSDNSNKKNAADLYLKP 240  
DB 181 GYVNPYTGRIRLDIOGTGQLLFVYINOLRLSDAQYICQAGDSDNSNKKNAADLYLKP 240  
QY 241 EPELYEDLRGSVTFHCALGPEVANYAKFLCROSSGENCDVYVNTLGKRAPAFEGRIILN 300  
DB 241 EPELYEDLRGSVTFHCALGPEVANYAKFLCROSSGENCDVYVNTLGKRAPAFEGRIILN 300  
QY 301 PODKDSFSVITGLRKEDAGRYLCAHSDDGQLQEGSPIOAMQLFVNEESTIPRSPTVVK 360  
DB 301 PODKDSFSVITGLRKEDAGRYLCAHSDDGQLQEGSPIOAMQLFVNEESTIPRSPTVVK 360  
QY 361 GVAAGSSVAVLCYPNRRESKSIKWCIMBGAONRCPLVDSEGMVKAQYEGRLSLEEPC 420  
DB 361 GVAAGSSVAVLCYPNRRESKSIKWCIMBGAONRCPLVDSEGMVKAQYEGRLSLEEPC 420  
QY 421 NGFTVILNOLTSRDGFWCLNGDTLMRTTVEIKIIEGEPNLKYPGNTAVLGETLKY 480  
DB 421 NGFTVILNOLTSRDGFWCLNGDTLMRTTVEIKIIEGEPNLKYPGNTAVLGETLKY 480  
QY 481 PCHFPCKFSSYEKYKWKNNNTGCOALPSODEGSKAFVNCDENSRLVSLTLNLYTRADEG 540  
DB 481 PCHFPCKFSSYEKYKWKNNNTGCOALPSODEGSKAFVNCDENSRLVSLTLNLYTRADEG 540  
QY 541 WYMGCVKQGHFYGETAAYVAVEERKAAGSDVSLAKADAAPDEKVLDSGFREIENKAIO 600  
DB 541 WYMGCVKQGHFYGETAAYVAVEERKAAGSDVSLAKADAAPDEKVLDSGFREIENKAIO 600  
QY 601 DPLFLAEEKAVADTRDOAGSRASVDSGSSSEEGGSRALVSTLVLGLVLAAGAAYGV 660  
DB 601 DPLFLAEEKAVADTRDOAGSRASVDSGSSSEEGGSRALVSTLVLGLVLAAGAAYGV 660  
QY 661 ARAHRKKNDRVSIIRSYRTDISMSDFENSREFGANDMGASSITQETSLGKKEEVAATTE 720

```

Db      661 AARHRKNDVRSISYRTDISMSDFENRREGADNNKASSITQETSLGKKEEFVATTE 720
OY      721 STTETKEPKKARRSSKEAEEMAYKDFLLQSSSTVAEAADGPOEA 764
Db      721 STTETKEPKKARRSSKEAEEMAYKDFLLQSSSTVAEAADGPOEA 764

RESULT 3
US-09-791-537-92050
; Sequence 92050, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Bionomix, Inc.
; APPLICANT: Dabe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBE
; FILE OF INVENTION: METHODS OF USE THEREOF
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 92050
; LENGTH: 764
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-791-537-92050

Query Match      100.0%: Score 764: DB 22: Length 764:
Best Local Similarity 100.0%: Pred. No. 0:
Matches 764: Conservative 0: Mismatches 0: Indels 0: Gaps 0:

OY      1 MLEFVTLCLAVFAIPSTKSPIFGPEEVNSVEGNSVSTICYPTSVNRHTRKRYCROGA 60
Db      1 MLEFVTLCLAVFAIPSTKSPIFGPEEVNSVEGNSVSTICYPTSVNRHTRKRYCROGA 60
OY      61 RGGCTTLLSSEGYSSSKYAGRANLTFNPENGTFVYVNIQLSODDSGRYKCGIGINSRGLS 120
Db      61 RGGCTTLLSSEGYSSSKYAGRANLTFNPENGTFVYVNIQLSODDSGRYKCGIGINSRGLS 120
OY      121 FDSVLEVSQGGPLNDTQVYTVDLGRVTYINCPEKTEAOKRSLYKQIGLYPVLYIDSS 180
Db      121 FDSVLEVSQGGPLNDTQVYTVDLGRVTYINCPEKTEAOKRSLYKQIGLYPVLYIDSS 180
OY      181 GYVNPNTYGRIRLDIOGTGOLLEFSVYINQLRLSDAGQYLCQAGDDSNKKNADQLVKP 240
Db      181 GYVNPNTYGRIRLDIOGTGOLLEFSVYINQLRLSDAGQYLCQAGDDSNKKNADQLVKP 240
OY      241 EPELVYEDLRSGSVTFHCLGPEVANVAKFLCROSSGECNDVYVNTLGKRAPAFEGRIILN 300
Db      241 EPELVYEDLRSGSVTFHCLGPEVANVAKFLCROSSGECNDVYVNTLGKRAPAFEGRIILN 300
OY      301 PODKDGFSVYITGLRKEDAGRYLCGAHSDGOLQEGSPFOAMQLFVNEESTIPRSPYVVK 360
Db      301 PODKDGFSVYITGLRKEDAGRYLCGAHSDGOLQEGSPFOAMQLFVNEESTIPRSPYVVK 360
OY      361 GVAAGSSVAVLCPPYRNKESKSIKYWCLMEGAONGRCPLLVDSGHWKAYEGRLSLLEBP 420
Db      361 GVAAGSSVAVLCPPYRNKESKSIKYWCLMEGAONGRCPLLVDSGHWKAYEGRLSLLEBP 420
OY      421 NGTFVILNQLTSRDAGRYWCLTNGDTLMRTTVEIKIIEGEPNLKVPGNVAVLGETLKV 480
Db      421 NGTFVILNQLTSRDAGRYWCLTNGDTLMRTTVEIKIIEGEPNLKVPGNVAVLGETLKV 480
OY      481 PCHEPCKFSSYEKXWCKNNNTGCOALPSODEGSPSAFVNCDENSRLVLTJNLVYRADEG 540
Db      481 PCHEPCKFSSYEKXWCKNNNTGCOALPSODEGSPSAFVNCDENSRLVLTJNLVYRADEG 540
OY      541 WYMGCVKQGHFYGETAAYVAVEERKAAGSRDVLAKADAAPEDEKVLDSGFEIENKAIQ 600
Db      541 WYMGCVKQGHFYGETAAYVAVEERKAAGSRDVLAKADAAPEDEKVLDSGFEIENKAIQ 600
OY      601 DPLFAEKAAYADTRDOAGDSRASVDSGSSSEBQGSRALVSTLVPLGLVLAAGVAVGV 660

```

```

Db      601 DPLFAEKAAYADTRDOAGDSRASVDSGSSSEBQGSRALVSTLVPLGLVLAAGVAVGV 660
OY      661 AARHRKNDVRSISYRTDISMSDFENRREGADNNKASSITQETSLGKKEEFVATTE 720
Db      661 AARHRKNDVRSISYRTDISMSDFENRREGADNNKASSITQETSLGKKEEFVATTE 720

RESULT 4
US-09-818-247-1
; Sequence 1, Application US/09818247
; GENERAL INFORMATION:
; APPLICANT: Mostov, Keith E.
; APPLICANT: Chapin, Steven J.
; APPLICANT: Richman-Eisenstat, Janice
; TITLE OF INVENTION: The Regents of the University of California
; TITLE OF INVENTION: Ligands Directed to the Non-Secretory Component,
; FILE REFERENCE: 18062E-000910US
; CURRENT APPLICATION NUMBER: US/09/818,247
; CURRENT FILING DATE: 2001-03-26
; PRIOR APPLICATION NUMBER: WO PCT/US01/09699
; PRIOR FILING DATE: 2001-03-26
; PRIOR FILING DATE: 2000-03-27
; PRIOR APPLICATION NUMBER: US 60/192,197
; PRIOR FILING DATE: 2000-03-27
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 764
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: human polymorphic immunoglobulin receptor (p1gR)
US-09-818-247-1

Query Match      100.0%: Score 764: DB 23: Length 764:
Best Local Similarity 100.0%: Pred. No. 0:
Matches 764: Conservative 0: Mismatches 0: Indels 0: Gaps 0:

OY      1 MLEFVTLCLAVFAIPSTKSPIFGPEEVNSVEGNSVSTICYPTSVNRHTRKRYCROGA 60
Db      1 MLEFVTLCLAVFAIPSTKSPIFGPEEVNSVEGNSVSTICYPTSVNRHTRKRYCROGA 60
OY      61 RGGCTTLLSSEGYSSSKYAGRANLTFNPENGTFVYVNIQLSODDSGRYKCGIGINSRGLS 120
Db      61 RGGCTTLLSSEGYSSSKYAGRANLTFNPENGTFVYVNIQLSODDSGRYKCGIGINSRGLS 120
OY      121 FDSVLEVSQGGPLNDTQVYTVDLGRVTYINCPEKTEAOKRSLYKQIGLYPVLYIDSS 180
Db      121 FDSVLEVSQGGPLNDTQVYTVDLGRVTYINCPEKTEAOKRSLYKQIGLYPVLYIDSS 180
OY      181 GYVNPNTYGRIRLDIOGTGOLLEFSVYINQLRLSDAGQYLCQAGDDSNKKNADQLVKP 240
Db      181 GYVNPNTYGRIRLDIOGTGOLLEFSVYINQLRLSDAGQYLCQAGDDSNKKNADQLVKP 240
OY      241 EPELVYEDLRSGSVTFHCLGPEVANVAKFLCROSSGECNDVYVNTLGKRAPAFEGRIILN 300
Db      241 EPELVYEDLRSGSVTFHCLGPEVANVAKFLCROSSGECNDVYVNTLGKRAPAFEGRIILN 300
OY      301 PODKDGFSVYITGLRKEDAGRYLCGAHSDGOLQEGSPFOAMQLFVNEESTIPRSPYVVK 360
Db      301 PODKDGFSVYITGLRKEDAGRYLCGAHSDGOLQEGSPFOAMQLFVNEESTIPRSPYVVK 360
OY      361 GVAAGSSVAVLCPPYRNKESKSIKYWCLMEGAONGRCPLLVDSGHWKAYEGRLSLLEBP 420
Db      361 GVAAGSSVAVLCPPYRNKESKSIKYWCLMEGAONGRCPLLVDSGHWKAYEGRLSLLEBP 420

```

OY 421 NGFTVILNOLTSRDAGFYWCLNNGDTLMTTVEIKIIEGEPNLKYPGNTAVLGETLKY 480  
DB 421 NGFTVILNOLTSRDAGFYWCLNNGDTLMTTVEIKIIEGEPNLKYPGNTAVLGETLKY 480  
OY 481 PCHEPKCFSSYEKWKMMNTGCOALPSODEGSKAFVNCDENSRLVSLTLNLTTRADEG 540  
DB 481 PCHEPKCFSSYEKWKMMNTGCOALPSODEGSKAFVNCDENSRLVSLTLNLTTRADEG 540  
OY 541 WYMGVYKOGHFYGETAAVYVAVEERKAAGSRDYSLAKADAAPEKVLDSGFREIENKAIO 600  
DB 541 WYMGVYKOGHFYGETAAVYVAVEERKAAGSRDYSLAKADAAPEKVLDSGFREIENKAIO 600  
OY 601 DPLFAEERKAVADTRQADGSRASVDSGSSSEOGSSRALVSTLVLGLVLAAGAVGV 660  
DB 601 DPLFAEERKAVADTRQADGSRASVDSGSSSEOGSSRALVSTLVLGLVLAAGAVGV 660  
OY 661 ARAHRKKNDRVIRSIRYRTDISMSDFENSREFGANDMGASSTIOETSLGKKEEFYATTE 720  
DB 661 ARAHRKKNDRVIRSIRYRTDISMSDFENSREFGANDMGASSTIOETSLGKKEEFYATTE 720  
OY 721 STTETKEPKKAKRSKEEMAYKDFLLQSSSTVAAEKODGPOEA 764  
DB 721 STTETKEPKKAKRSKEEMAYKDFLLQSSSTVAAEKODGPOEA 764

## RESULT 5

US-09-949-016-6254  
; Sequence 6254, Application US/09949016  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949, 016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: fastseq for Windows Version 4.0  
; SEQ ID NO 6254  
; LENGTH: 764  
; TYPE: PRT  
; ORGANISM: Human  
US-09-949-016-6254

Query Match 100.0%; Score 764; DB 24; Length 764;

Best Local Similarity 100.0%; Pred. No. 0; Mismatches 0; Indels 0; Gaps 0;

Matches 764; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 MLFVLTLCLAVEPAISTKSPIFGPEEVNSVEGNSVITCYPTSVNRRHTRKYMROGA 60  
DB 1 MLFVLTLCLAVEPAISTKSPIFGPEEVNSVEGNSVITCYPTSVNRRHTRKYMROGA 60  
OY 61 RGCITLISSEGYVSSKYAGRANLJNFPENGTFVVAIAQLSODDSGRYKCGLGINSRGLS 120  
DB 61 RGCITLISSEGYVSSKYAGRANLJNFPENGTFVVAIAQLSODDSGRYKCGLGINSRGLS 120  
OY 121 FDVLSLEVSOGPGLNLTNTKYTVDLGRTVTINCPFKTENAKRKSLYKQIGLFPVLYIDSS 180  
DB 121 FDVLSLEVSOGPGLNLTNTKYTVDLGRTVTINCPFKTENAKRKSLYKQIGLFPVLYIDSS 180  
OY 181 GYVNPNTYGRIRLDIOGTGOLFSVYINOLRLSDAGQYLCQAGDDSNMKNKNDLQVLRP 240  
DB 181 GYVNPNTYGRIRLDIOGTGOLFSVYINOLRLSDAGQYLCQAGDDSNMKNKNDLQVLRP 240  
OY 241 EPELVYEDLRGSYTFHCALGPEVANAANKFLCROSSGSCNCDVYVNTLGRAPAFEGRIILN 300  
DB 241 EPELVYEDLRGSYTFHCALGPEVANAANKFLCROSSGSCNCDVYVNTLGRAPAFEGRIILN 300

OY 301 PODKGSFSFVITGLRKEAGRYLCAHSQDQLESPIQANOLFVNEESTIPRSPYVK 360  
DB 301 PODKGSFSFVITGLRKEAGRYLCAHSQDQLESPIQANOLFVNEESTIPRSPYVK 360  
OY 361 GYVSSVAVLCPYNNRESKSIKWKCMLEGAQONRCPLVLDSEGVAAQYEGRLSLLEERG 420  
DB 361 GYVSSVAVLCPYNNRESKSIKWKCMLEGAQONRCPLVLDSEGVAAQYEGRLSLLEERG 420  
OY 421 NGFTVILNOLTSRDAGFYWCLNNGDTLMTTVEIKIIEGEPNLKYPGNTAVLGETLKY 480  
DB 421 NGFTVILNOLTSRDAGFYWCLNNGDTLMTTVEIKIIEGEPNLKYPGNTAVLGETLKY 480  
OY 481 PCHEPKCFSSYEKWKMMNTGCOALPSODEGSKAFVNCDENSRLVSLTLNLTTRADEG 540  
DB 481 PCHEPKCFSSYEKWKMMNTGCOALPSODEGSKAFVNCDENSRLVSLTLNLTTRADEG 540  
OY 541 WYMGVYKOGHFYGETAAVYVAVEERKAAGSRDYSLAKADAAPEKVLDSGFREIENKAIO 600  
DB 541 WYMGVYKOGHFYGETAAVYVAVEERKAAGSRDYSLAKADAAPEKVLDSGFREIENKAIO 600  
OY 601 DPLFAEERKAVADTRQADGSRASVDSGSSSEOGSSRALVSTLVLGLVLAAGAVGV 660  
DB 601 DPLFAEERKAVADTRQADGSRASVDSGSSSEOGSSRALVSTLVLGLVLAAGAVGV 660  
OY 661 ARAHRKKNDRVIRSIRYRTDISMSDFENSREFGANDMGASSTIOETSLGKKEEFYATTE 720  
DB 661 ARAHRKKNDRVIRSIRYRTDISMSDFENSREFGANDMGASSTIOETSLGKKEEFYATTE 720  
OY 721 STTETKEPKKAKRSKEEMAYKDFLLQSSSTVAAEKODGPOEA 764  
DB 721 STTETKEPKKAKRSKEEMAYKDFLLQSSSTVAAEKODGPOEA 764

## RESULT 6

US-09-949-039-65  
; Sequence 65, Application US/09949039  
; GENERAL INFORMATION:  
; APPLICANT: HAWLEY, STEPHEN B.  
; TITLE OF INVENTION: COMPOUNDS AND MOLECULAR COMPLEXES COMPRISING MULTIPLE  
; TITLE OF INVENTION: BINDING REGIONS DIRECTED TO TRANSCRIPTOTIC LIGANDS  
; FILE REFERENCE: 057220/1301  
; CURRENT APPLICATION NUMBER: US/09/949, 039  
; CURRENT FILING DATE: 2001-09-06  
; NUMBER OF SEQ ID NOS: 114  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 65  
; LENGTH: 764  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-949-039-65

Query Match 100.0%; Score 764; DB 24; Length 764;

Best Local Similarity 100.0%; Pred. No. 0; Mismatches 0; Indels 0; Gaps 0;

Matches 764; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 MLFVLTLCLAVEPAISTKSPIFGPEEVNSVEGNSVITCYPTSVNRRHTRKYMROGA 60  
DB 1 MLFVLTLCLAVEPAISTKSPIFGPEEVNSVEGNSVITCYPTSVNRRHTRKYMROGA 60  
OY 61 RGCITLISSEGYVSSKYAGRANLJNFPENGTFVVAIAQLSODDSGRYKCGLGINSRGLS 120  
DB 61 RGCITLISSEGYVSSKYAGRANLJNFPENGTFVVAIAQLSODDSGRYKCGLGINSRGLS 120  
OY 121 FDVLSLEVSOGPGLNLTNTKYTVDLGRTVTINCPFKTENAKRKSLYKQIGLFPVLYIDSS 180  
DB 121 FDVLSLEVSOGPGLNLTNTKYTVDLGRTVTINCPFKTENAKRKSLYKQIGLFPVLYIDSS 180  
OY 181 GYVNPNTYGRIRLDIOGTGOLFSVYINOLRLSDAGQYLCQAGDDSNMKNKNDLQVLRP 240  
DB 181 GYVNPNTYGRIRLDIOGTGOLFSVYINOLRLSDAGQYLCQAGDDSNMKNKNDLQVLRP 240  
OY 241 EPELVYEDLRGSYTFHCALGPEVANAANKFLCROSSGSCNCDVYVNTLGRAPAFEGRIILN 300  
DB 241 EPELVYEDLRGSYTFHCALGPEVANAANKFLCROSSGSCNCDVYVNTLGRAPAFEGRIILN 300

Db 241 EPELVYEDLRGSVTFHCLGPEVANNAKFLCROSSGENCDVYVNTLGRAPAFEGRIILN 300  
Qy 301 PQRDGSFSVYITGLRKEDAGRYLCAHSDGLOEGSP IQAMOLFVNEESTIPRSPYVK 360  
Db 301 PQRDGSFSVYITGLRKEDAGRYLCAHSDGLOEGSP IQAMOLFVNEESTIPRSPYVK 360  
Qy 361 GVAGSSVAVLCPRYNRKESKSIKYWCLMEGAONGRCPLLYDSEGVVKAQYEGRLSLEBPG 420  
Db 361 GVAGSSVAVLCPRYNRKESKSIKYWCLMEGAONGRCPLLYDSEGVVKAQYEGRLSLEBPG 420  
Qy 421 NGFTFVLINOLTSRDAGRYWCLTNGDITMRTTVEIKIIEGPNLKVGNNTAVAGETLKV 480  
Db 421 NGFTFVLINOLTSRDAGRYWCLTNGDITMRTTVEIKIIEGPNLKVGNNTAVAGETLKV 480  
Qy 481 PCHFPCKFSSEYKWKNNNGCOALPSODEGPKAFVNCDENSRLVSLTLNLTVRADeg 540  
Db 481 PCHFPCKFSSEYKWKNNNGCOALPSODEGPKAFVNCDENSRLVSLTLNLTVRADeg 540  
Qy 541 WYMGVKGHFGYGETAAVYVAVERKKAAGSDVSLAKADAAPEKVLDSGFRLEENKAIQ 600  
Db 541 WYMGVKGHFGYGETAAVYVAVERKKAAGSDVSLAKADAAPEKVLDSGFRLEENKAIQ 600  
Qy 601 DPLFAEKAADTRDQADGSRASVDSGSSSEOGSSRALVSTLPLGLVLAAGVAVGV 660  
Db 601 DPLFAEKAADTRDQADGSRASVDSGSSSEOGSSRALVSTLPLGLVLAAGVAVGV 660  
Qy 661 ARAHRKKNVDVRSIRSYRTDISMSDFENSREFGANDNMGASSITQETSLGKEEFVATTE 720  
Db 661 ARAHRKKNVDVRSIRSYRTDISMSDFENSREFGANDNMGASSITQETSLGKEEFVATTE 720  
Qy 721 STETKEPKKAKRSSKEEAEMAYKDFLLQSTVAEAADGPOEA 764  
Db 721 STETKEPKKAKRSSKEEAEMAYKDFLLQSTVAEAADGPOEA 764

RESULT 7  
US-09-981-353-59  
: Sequence 59, Application US/09981353  
: GENERAL INFORMATION:  
: APPLICANT: Lasek, Amy W.  
: APPLICANT: Jones, David A.  
: TITLE OF INVENTION: GENES EXPRESSED IN COLON CANCER  
: FILE REFERENCE: PA-0038 US  
: CURRENT APPLICATION NUMBER: US/09/981,353  
: CURRENT FILING DATE: 2001-10-11  
: NUMBER OF SEQ ID NOS: 194  
: SOFTWARE: PERL Program  
: SEQ ID NO 59  
: LENGTH: 764  
: TYPE: PRT  
: ORGANISM: Homo sapiens  
: FEATURE:  
: NAME/KEY: misc\_feature  
: OTHER INFORMATION: Incyte ID No: 2680109CD1  
US-09-981-353-59

Query Match 100.0%; Score 764; DB 25; Length 764;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 764; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MLEFVLTCLAVPAFAISTKSPIFGPEEVNSVEGNSVSTCYPTPTSVNRHTRKRYWCROGA 60  
Db 1 MLEFVLTCLAVPAFAISTKSPIFGPEEVNSVEGNSVSTCYPTPTSVNRHTRKRYWCROGA 60  
Qy 61 RGGCTTLISSEGYVSSKAGRANLTNPENCTFVYVNTAQLSDODSGRKCGLGINSRGLS 120  
Db 61 RGGCTTLISSEGYVSSKAGRANLTNPENCTFVYVNTAQLSDODSGRKCGLGINSRGLS 120  
Qy 121 FDVSLSEVSGPGLNDTRKYTVTDLGRVTYINCPEFTEAOKRKSILYKQIGLYPVLIDSS 180  
Db 121 FDVSLSEVSGPGLNDTRKYTVTDLGRVTYINCPEFTEAOKRKSILYKQIGLYPVLIDSS 180  
Qy 181 GYVNPNTYGRIRLDIOGTGLFSVYINQLRLSDAGGYLCOAGDDSNSNKNADLOVLKP 240

Db 181 GYVNPNTYGRIRLDIOGTGLFSVYINQLRLSDAGGYLCOAGDDSNSNKNADLOVLKP 240  
Qy 241 EPELVYEDLRGSVTFHCLGPEVANNAKFLCROSSGENCDVYVNTLGRAPAFEGRIILN 300  
Db 241 EPELVYEDLRGSVTFHCLGPEVANNAKFLCROSSGENCDVYVNTLGRAPAFEGRIILN 300  
Qy 301 PQRDGSFSVYITGLRKEDAGRYLCAHSDGLOEGSP IQAMOLFVNEESTIPRSPYVK 360  
Db 301 PQRDGSFSVYITGLRKEDAGRYLCAHSDGLOEGSP IQAMOLFVNEESTIPRSPYVK 360  
Qy 361 GVAGSSVAVLCPRYNRKESKSIKYWCLMEGAONGRCPLLYDSEGVVKAQYEGRLSLEBPG 420  
Db 361 GVAGSSVAVLCPRYNRKESKSIKYWCLMEGAONGRCPLLYDSEGVVKAQYEGRLSLEBPG 420  
Qy 421 NGFTFVLINOLTSRDAGRYWCLTNGDITMRTTVEIKIIEGPNLKVGNNTAVAGETLKV 480  
Db 421 NGFTFVLINOLTSRDAGRYWCLTNGDITMRTTVEIKIIEGPNLKVGNNTAVAGETLKV 480  
Qy 481 PCHFPCKFSSEYKWKNNNGCOALPSODEGPKAFVNCDENSRLVSLTLNLTVRADeg 540  
Db 481 PCHFPCKFSSEYKWKNNNGCOALPSODEGPKAFVNCDENSRLVSLTLNLTVRADeg 540  
Qy 541 WYMGVKGHFGYGETAAVYVAVERKKAAGSDVSLAKADAAPEKVLDSGFRLEENKAIQ 600  
Db 541 WYMGVKGHFGYGETAAVYVAVERKKAAGSDVSLAKADAAPEKVLDSGFRLEENKAIQ 600  
Qy 601 DPLFAEKAADTRDQADGSRASVDSGSSSEOGSSRALVSTLPLGLVLAAGVAVGV 660  
Db 601 DPLFAEKAADTRDQADGSRASVDSGSSSEOGSSRALVSTLPLGLVLAAGVAVGV 660  
Qy 661 ARAHRKKNVDVRSIRSYRTDISMSDFENSREFGANDNMGASSITQETSLGKEEFVATTE 720  
Db 661 ARAHRKKNVDVRSIRSYRTDISMSDFENSREFGANDNMGASSITQETSLGKEEFVATTE 720  
Qy 721 STETKEPKKAKRSSKEEAEMAYKDFLLQSTVAEAADGPOEA 764  
Db 721 STETKEPKKAKRSSKEEAEMAYKDFLLQSTVAEAADGPOEA 764

RESULT 8  
US-09-989-919-124  
: Sequence 124, Application US/09989919  
: GENERAL INFORMATION:  
: APPLICANT: Macina, Roberto  
: APPLICANT: Recipon, Herve  
: APPLICANT: Pluta, Jason  
: APPLICANT: Ghosh, Malaivika  
: APPLICANT: Liu, Chenghua  
: APPLICANT: Sun, Yongming  
: TITLE OF INVENTION: Compositions and Methods Relating to Colon Specific Genes  
: FILE REFERENCE: DEX-0289  
: CURRENT APPLICATION NUMBER: US/09/989,919  
: CURRENT FILING DATE: 2001-11-21  
: PRIOR APPLICATION NUMBER: 60/252,505  
: PRIOR FILING DATE: 2000-11-22  
: NUMBER OF SEQ ID NOS: 124  
: SOFTWARE: PatentIn version 3.1  
: SEQ ID NO 124  
: LENGTH: 764  
: TYPE: PRT  
: ORGANISM: Homo sapien  
US-09-989-919-124

Query Match 100.0%; Score 764; DB 25; Length 764;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 764; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MLEFVLTCLAVPAFAISTKSPIFGPEEVNSVEGNSVSTCYPTPTSVNRHTRKRYWCROGA 60  
Db 1 MLEFVLTCLAVPAFAISTKSPIFGPEEVNSVEGNSVSTCYPTPTSVNRHTRKRYWCROGA 60  
Qy 61 RGGCTTLISSEGYVSSKAGRANLTNPENCTFVYVNTAQLSDODSGRKCGLGINSRGLS 120

```

Db 61 RGGCITLISSEGVSSKYAGRAUJTFNPENGTFVVAIAQLSODDSGRYKCGILINSRGLS 120
Qy 121 FDVLSLEVSGPGLNDTKYTYVDLGRVTIINCPFKTENAKRSLXKQIGLFPVLVIDSS 180
Db 121 FDVLSLEVSGPGLNDTKYTYVDLGRVTIINCPFKTENAKRSLXKQIGLFPVLVIDSS 180
Qy 181 GYVNPYTGRIRLDIOGTGQLFSSVYINQLRLSDAQOYLCOAGDDSNKKNADLOVLRP 240
Db 181 GYVNPYTGRIRLDIOGTGQLFSSVYINQLRLSDAQOYLCOAGDDSNKKNADLOVLRP 240
Qy 241 EPELYVEDLRGSVTFHCALGPEVANAFAKFLCROSSGSCNCDVYNTLGKRAPAFEGRIILN 300
Db 241 EPELYVEDLRGSVTFHCALGPEVANAFAKFLCROSSGSCNCDVYNTLGKRAPAFEGRIILN 300
Qy 301 PODKDSFSVYITGLRKEDAGRYLCAHSDGLOEGSP1QAMOLFVNEEST1PRSP1TVK 360
Db 301 PODKDSFSVYITGLRKEDAGRYLCAHSDGLOEGSP1QAMOLFVNEEST1PRSP1TVK 360
Qy 361 GYAGSSVAVLCPYNRKESKSIKYWCLMEGAONGRCPLVDSEGVNAQYEGRLSLLEBPG 420
Db 361 GYAGSSVAVLCPYNRKESKSIKYWCLMEGAONGRCPLVDSEGVNAQYEGRLSLLEBPG 420
Qy 421 NGFTVILNOLTSRDAGFYWCLTNGDTLMRTTVEIKIIEGEPNLKYPGNTAVLGETLKV 480
Db 421 NGFTVILNOLTSRDAGFYWCLTNGDTLMRTTVEIKIIEGEPNLKYPGNTAVLGETLKV 480
Qy 481 PCHEPKCFSSYETKWKMNNTGCOALPSODEGSKAFVNCDENSRLVSLTLNLTTRADEG 540
Db 481 PCHEPKCFSSYETKWKMNNTGCOALPSODEGSKAFVNCDENSRLVSLTLNLTTRADEG 540
Qy 541 WYMGVKGQHFYGETAAYVAAVEERKAAGSRDVS1AKADAAPDEKVLDSGFREIENKA1Q 600
Db 541 WYMGVKGQHFYGETAAYVAAVEERKAAGSRDVS1AKADAAPDEKVLDSGFREIENKA1Q 600
Qy 601 DPLFLAEKAVADTRDOADGSRASVDSGSSSEOGSSRALVSTLVLPLGLVLAAGAVAVG 660
Db 601 DPLFLAEKAVADTRDOADGSRASVDSGSSSEOGSSRALVSTLVLPLGLVLAAGAVAVG 660
Qy 661 ARAHRRKNDVRS1RSYRTDISMSDFENSRFEGANDMGASST1QETSLGKKEEFVATTE 720
Db 661 ARAHRRKNDVRS1RSYRTDISMSDFENSRFEGANDMGASST1QETSLGKKEEFVATTE 720
Qy 721 STTETKEPKKAKRSSKEEAMAYKDFLQSTVAEAADGPOEA 764
Db 721 STTETKEPKKAKRSSKEEAMAYKDFLQSTVAEAADGPOEA 764

```

## RESULT 9

```

US-60-239-841-59
: Sequence 59, Application US/60239841
: GENERAL INFORMATION:
: APPLICANT: Lasek, Amy W.
: TITLE OF INVENTION: GENES EXPRESSED IN COLON CANCER
: FILE REFERENCE: PA-0038 P
: CURRENT APPLICATION NUMBER: US/60/239, 841
: NUMBER OF SEQ ID NOS: 194
: SOFTWARE: PERL Program
: SEQ ID NO 59
: LENGTH: 764
: TYPE: PRT
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: misc.feature
: OTHER INFORMATION: Incyte ID No: 2680109CD1
US-60-239-841-59

```

```

Query Match 100.0%; Score 764; DB 32; Length 764;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 764; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy 1 MLFLVTLCLLAVPAISTKSPIEGPEVANSVEGNSSTICCYPTPTSVNHRTRKMYCROGA 60
Db 1 MLFLVTLCLLAVPAISTKSPIEGPEVANSVEGNSSTICCYPTPTSVNHRTRKMYCROGA 60
Qy 61 RGGCITLISSEGVSSKYAGRAUJTFNPENGTFVVAIAQLSODDSGRYKCGILINSRGLS 120
Db 61 RGGCITLISSEGVSSKYAGRAUJTFNPENGTFVVAIAQLSODDSGRYKCGILINSRGLS 120
Qy 121 FDVLSLEVSGPGLNDTKYTYVDLGRVTIINCPFKTENAKRSLXKQIGLFPVLVIDSS 180
Db 121 FDVLSLEVSGPGLNDTKYTYVDLGRVTIINCPFKTENAKRSLXKQIGLFPVLVIDSS 180
Qy 181 GYVNPYTGRIRLDIOGTGQLFSSVYINQLRLSDAQOYLCOAGDDSNKKNADLOVLRP 240
Db 181 GYVNPYTGRIRLDIOGTGQLFSSVYINQLRLSDAQOYLCOAGDDSNKKNADLOVLRP 240
Qy 241 EPELYVEDLRGSVTFHCALGPEVANAFAKFLCROSSGSCNCDVYNTLGKRAPAFEGRIILN 300
Db 241 EPELYVEDLRGSVTFHCALGPEVANAFAKFLCROSSGSCNCDVYNTLGKRAPAFEGRIILN 300
Qy 301 PODKDSFSVYITGLRKEDAGRYLCAHSDGLOEGSP1QAMOLFVNEEST1PRSP1TVK 360
Db 301 PODKDSFSVYITGLRKEDAGRYLCAHSDGLOEGSP1QAMOLFVNEEST1PRSP1TVK 360
Qy 361 GYAGSSVAVLCPYNRKESKSIKYWCLMEGAONGRCPLVDSEGVNAQYEGRLSLLEBPG 420
Db 361 GYAGSSVAVLCPYNRKESKSIKYWCLMEGAONGRCPLVDSEGVNAQYEGRLSLLEBPG 420
Qy 421 NGFTVILNOLTSRDAGFYWCLTNGDTLMRTTVEIKIIEGEPNLKYPGNTAVLGETLKV 480
Db 421 NGFTVILNOLTSRDAGFYWCLTNGDTLMRTTVEIKIIEGEPNLKYPGNTAVLGETLKV 480
Qy 481 PCHEPKCFSSYETKWKMNNTGCOALPSODEGSKAFVNCDENSRLVSLTLNLTTRADEG 540
Db 481 PCHEPKCFSSYETKWKMNNTGCOALPSODEGSKAFVNCDENSRLVSLTLNLTTRADEG 540
Qy 541 WYMGVKGQHFYGETAAYVAAVEERKAAGSRDVS1AKADAAPDEKVLDSGFREIENKA1Q 600
Db 541 WYMGVKGQHFYGETAAYVAAVEERKAAGSRDVS1AKADAAPDEKVLDSGFREIENKA1Q 600
Qy 601 DPLFLAEKAVADTRDOADGSRASVDSGSSSEOGSSRALVSTLVLPLGLVLAAGAVAVG 660
Db 601 DPLFLAEKAVADTRDOADGSRASVDSGSSSEOGSSRALVSTLVLPLGLVLAAGAVAVG 660
Qy 661 ARAHRRKNDVRS1RSYRTDISMSDFENSRFEGANDMGASST1QETSLGKKEEFVATTE 720
Db 661 ARAHRRKNDVRS1RSYRTDISMSDFENSRFEGANDMGASST1QETSLGKKEEFVATTE 720
Qy 721 STTETKEPKKAKRSSKEEAMAYKDFLQSTVAEAADGPOEA 764
Db 721 STTETKEPKKAKRSSKEEAMAYKDFLQSTVAEAADGPOEA 764

```

## RESULT 10

```

US-09-717-888-4
: Sequence 4, Application US/09717888
: GENERAL INFORMATION:
: APPLICANT: ANDREW C. HIATT, JULIAN
: TITLE OF INVENTION: IMMUNOGLOBULINS CONTAINING PROTECTION
: NUMBER OF SEQUENCES: 19
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Lyon & Lyon
: STREET: 633 West Fifth Street
: Suite 4700
: CITY: Los Angeles
: STATE: California
: COUNTRY: U.S.A.
: ZIP: 90071
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
storage

```

```
COMPUTER: IBM Compatible
OPERATING SYSTEM: IBM P.C. DOS 5.0
SOFTWARE: Word Perfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/717.888
FILING DATE: 20-Nov-2000
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/434.000
FILING DATE: <Unknown>
APPLICATION NUMBER: 08/367.395
FILING DATE: 30-Dec-94
ATTORNEY/AGENT INFORMATION:
NAME: Guise, Jeffrey W.
REGISTRATION NUMBER: 34,613
REFERENCE/DOCKET NUMBER: 212/127
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 552-8400
TELEFAX: (619) 552-0159
TELEX: 67-3510
SEQUENCE LISTING
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 746 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
DESCRIPTION: Human Polymunoglobulin Receptor
US-09-717-888-4
Query Match          97.6%; Score 746; DB 21; Length 746;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 746; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 19 KSPFEGEEVNSVSGNSVITCYPTPTSVNRHTRKRYKRCGARGCITLLISSEGYSSKY 78
DB 1 KSPFEGEEVNSVSGNSVITCYPTPTSVNRHTRKRYKRCGARGCITLLISSEGYSSKY 60
QY 79 AGRANLTNFPENGFFVNNIAQLSODDSGRYKCGIGINSRGLSPDVSLEVSOGPLNDTK 138
DB 61 AGRANLTNFPENGFFVNNIAQLSODDSGRYKCGIGINSRGLSPDVSLEVSOGPLNDTK 120
QY 139 VYTVDLGRVTYINCPEFTENAOKRSLXKQIGLYPVLVIDSSGYVNPNTGRIRLDIQGT 198
DB 121 VYTVDLGRVTYINCPEFTENAOKRSLXKQIGLYPVLVIDSSGYVNPNTGRIRLDIQGT 180
QY 199 GQLLFVYINQLRLSDAGQYLCQAGDSDNSNKKNAADQVLPPELVEYEDLRGSVTFHCA 258
DB 181 GQLLFVYINQLRLSDAGQYLCQAGDSDNSNKKNAADQVLPPELVEYEDLRGSVTFHCA 240
QY 259 LGPEVANVAFAFLCRSGSGENCDDVYVNTLGRAPAFEGRIILNPDOKDGSFVYITGLRKE 318
DB 241 LGPEVANVAFAFLCRSGSGENCDDVYVNTLGRAPAFEGRIILNPDOKDGSFVYITGLRKE 300
QY 319 DAGRYLCGASHDGLQSGSPIQAMQLFVNEESTIPRSPYVKGAGSSVAVLCPYNRKES 378
DB 301 DAGRYLCGASHDGLQSGSPIQAMQLFVNEESTIPRSPYVKGAGSSVAVLCPYNRKES 360
QY 379 KSIRYKWLMEGAONGRCPLVDSGKWKAOYEGRLSLLEPGNGTFTVILNQLTSRDAGF 438
DB 361 KSIRYKWLMEGAONGRCPLVDSGKWKAOYEGRLSLLEPGNGTFTVILNQLTSRDAGF 420
QY 439 YWCLTNGDTLMRTTVEIKIIEGEPNLKPGNVTAVLGETLKVPCHPFCKSSYKRYCKW 498
DB 421 YWCLTNGDTLMRTTVEIKIIEGEPNLKPGNVTAVLGETLKVPCHPFCKSSYKRYCKW 480
QY 499 NNTGCOALPSODEGSPSAFYVNCDENSRLVSLTLNLVYRADEGMYWCGVKQGHFYGETAAV 558
DB 481 NNTGCOALPSODEGSPSAFYVNCDENSRLVSLTLNLVYRADEGMYWCGVKQGHFYGETAAV 540
QY 559 YVAVEERKAAGSRVSLAKADAARDEKVLDSGFEIENKAIODPRLFAEERKAAVADTRDOA 618
DB 541 YVAVEERKAAGSRVSLAKADAARDEKVLDSGFEIENKAIODPRLFAEERKAAVADTRDOA 600
```

```
541 YVAVEERKAAGSRVSLAKADAARDEKVLDSGFEIENKAIODPRLFAEERKAAVADTRDOA 600
QY 619 DGSRAVSDGSSSEOGGSSRALVSTVPLGLVLAAGAAGVARRRRKRVDRYSIRSYR 678
DB 601 DGSRAVSDGSSSEOGGSSRALVSTVPLGLVLAAGAAGVARRRRKRVDRYSIRSYR 660
QY 679 TDISMSPFENSREGANDNNKASSITQETSLGKEEVAATTESTETKEKKAKKRSKEE 738
DB 661 TDISMSPFENSREGANDNNKASSITQETSLGKEEVAATTESTETKEKKAKKRSKEE 720
QY 739 AEMAYKDFLQSSSTVAEAADGPQEA 764
DB 721 AEMAYKDFLQSSSTVAEAADGPQEA 746
RESULT 11
US-09-982-107-4
Sequence 4: Application US/09982107
GENERAL INFORMATION:
APPLICANT: HIATT, ANDREW C.
APPLICANT: HEIN, MICHAEL B.
TITLE OF INVENTION: METHODS FOR PRODUCING IMMUNOGLOBULINS CONTAINING
TITLE OF INVENTION: PROTECTION PROTEINS IN PLANTS AND THEIR USE
FILE REFERENCE: EP13002E
CURRENT APPLICATION NUMBER: US/09/982.107
CURRENT FILING DATE: 2001-10-16
NUMBER OF SEQ ID NOS: 19
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO: 4
LENGTH: 746
TYPE: PRT
ORGANISM: Homo sapiens
US-09-982-107-4
Query Match          97.6%; Score 746; DB 25; Length 746;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 746; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 19 KSPFEGEEVNSVSGNSVITCYPTPTSVNRHTRKRYKRCGARGCITLLISSEGYSSKY 78
DB 1 KSPFEGEEVNSVSGNSVITCYPTPTSVNRHTRKRYKRCGARGCITLLISSEGYSSKY 60
QY 79 AGRANLTNFPENGFFVNNIAQLSODDSGRYKCGIGINSRGLSPDVSLEVSOGPLNDTK 138
DB 61 AGRANLTNFPENGFFVNNIAQLSODDSGRYKCGIGINSRGLSPDVSLEVSOGPLNDTK 120
QY 139 VYTVDLGRVTYINCPEFTENAOKRSLXKQIGLYPVLVIDSSGYVNPNTGRIRLDIQGT 198
DB 121 VYTVDLGRVTYINCPEFTENAOKRSLXKQIGLYPVLVIDSSGYVNPNTGRIRLDIQGT 180
QY 199 GQLLFVYINQLRLSDAGQYLCQAGDSDNSNKKNAADQVLPPELVEYEDLRGSVTFHCA 258
DB 181 GQLLFVYINQLRLSDAGQYLCQAGDSDNSNKKNAADQVLPPELVEYEDLRGSVTFHCA 240
QY 259 LGPEVANVAFAFLCRSGSGENCDDVYVNTLGRAPAFEGRIILNPDOKDGSFVYITGLRKE 318
DB 241 LGPEVANVAFAFLCRSGSGENCDDVYVNTLGRAPAFEGRIILNPDOKDGSFVYITGLRKE 300
QY 319 DAGRYLCGASHDGLQSGSPIQAMQLFVNEESTIPRSPYVKGAGSSVAVLCPYNRKES 378
DB 301 DAGRYLCGASHDGLQSGSPIQAMQLFVNEESTIPRSPYVKGAGSSVAVLCPYNRKES 360
QY 379 KSIRYKWLMEGAONGRCPLVDSGKWKAOYEGRLSLLEPGNGTFTVILNQLTSRDAGF 438
DB 361 KSIRYKWLMEGAONGRCPLVDSGKWKAOYEGRLSLLEPGNGTFTVILNQLTSRDAGF 420
QY 439 YWCLTNGDTLMRTTVEIKIIEGEPNLKPGNVTAVLGETLKVPCHPFCKSSYKRYCKW 498
DB 421 YWCLTNGDTLMRTTVEIKIIEGEPNLKPGNVTAVLGETLKVPCHPFCKSSYKRYCKW 480
QY 499 NNTGCOALPSODEGSPSAFYVNCDENSRLVSLTLNLVYRADEGMYWCGVKQGHFYGETAAV 558
DB 481 NNTGCOALPSODEGSPSAFYVNCDENSRLVSLTLNLVYRADEGMYWCGVKQGHFYGETAAV 540
```

QY 559 YVAVERKAAGSRVSLAKADAAPEKVDLSGFEIENKAIODPRLEAEKAVADTRDQA 618  
DB 541 YVAVERKAAGSRVSLAKADAAPEKVDLSGFEIENKAIODPRLEAEKAVADTRDQA 600  
QY 619 DGSASVDSGSSSEOGSSRALVSTVPLGLVLAAGVAVARARRKRVDRYSIRSYR 678  
DB 601 DGSASVDSGSSSEOGSSRALVSTVPLGLVLAAGVAVARARRKRVDRYSIRSYR 660  
QY 679 TDISMDSFENSREFGANDNMKASSITQETSLGKKEEVPATTESTETTEKPKAKRSSKEE 738  
DB 661 TDISMDSFENSREFGANDNMKASSITQETSLGKKEEVPATTESTETTEKPKAKRSSKEE 720  
QY 739 AEMAYKDFLLQSSVVAEADQDPOEA 764  
DB 721 AEMAYKDFLLQSSVVAEADQDPOEA 746

## RESULT 12

US-60-443-566-4535  
Sequence 4535, Application US/60443566  
GENERAL INFORMATION:  
APPLICANT: CARGILL, Michele  
APPLICANT: BEGOVICH, Ann  
TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH  
TITLE OF INVENTION: RHEUMATOID ARTHRITIS, METHODS OF DETECTION AND USES THEREOF  
FILE REFERENCE: CL001447  
CURRENT APPLICATION NUMBER: US/60/443, 566  
CURRENT FILING DATE: 2003-01-30  
NUMBER OF SEQ ID NOS: 25102  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 4535  
LENGTH: 764  
TYPE: PRF  
ORGANISM: Homo sapiens  
US-60-443-566-4535

Query Match 86.8%; Score 663; DB 32; Length 764;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 763; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MLFVLTCLAVPAISTKSPIFGPEEVNSVEGNSVITCYPTSVNHRTRKRCROGA 60  
DB 1 MLFVLTCLAVPAISTKSPIFGPEEVNSVEGNSVITCYPTSVNHRTRKRCROGA 60  
QY 61 RGCITLISSEGVSSKYAGRANLTFNPNENGTFVYVNIQLOSDSGRYKCGLGINSRGLS 120  
DB 61 RGCITLISSEGVSSKYAGRANLTFNPNENGTFVYVNIQLOSDSGRYKCGLGINSRGLS 120  
QY 121 FVLSLEVSOGPGLNDTKYTVDLGRTVITNCPFKTENAKRKSILYKQIGLPLVYIDS 180  
DB 121 FVLSLEVSOGPGLNDTKYTVDLGRTVITNCPFKTENAKRKSILYKQIGLPLVYIDS 180  
QY 181 GYVNPYNGRIRLDIOGTGQLFSVYINQLRLSDAGQYLCQAGDSDNSNKKNDLOVLPK 240  
DB 181 GYVNPYNGRIRLDIOGTGQLFSVYINQLRLSDAGQYLCQAGDSDNSNKKNDLOVLPK 240  
QY 241 EPELVYEDLRGSVTFHCALGPEVANYAKFLCROSSGENDVYVNLGKRAPAFEGRIILN 300  
DB 241 EPELVYEDLRGSVTFHCALGPEVANYAKFLCROSSGENDVYVNLGKRAPAFEGRIILN 300  
QY 301 POKKDSFSVYITGLRKEDAGRYLCGAHSDGLOEGSPIOAMOLFVNEESTIPRSPTVYK 360  
DB 301 POKKDSFSVYITGLRKEDAGRYLCGAHSDGLOEGSPIOAMOLFVNEESTIPRSPTVYK 360  
QY 361 GYAGSSVAVLCPYNRKESKSIKYWCLMEGAONGRCPPLVDSGVAQYEGRLSLEBPG 420  
DB 361 GYAGSSVAVLCPYNRKESKSIKYWCLMEGAONGRCPPLVDSGVAQYEGRLSLEBPG 420  
QY 421 NGFTFYILNQLSRDAGFYWCLTNGDTLMRTTVEIKIIEGEPNLKVPNGVTAVLGFTLVK 480  
DB 421 NGFTFYILNQLSRDAGFYWCLTNGDTLMRTTVEIKIIEGEPNLKVPNGVTAVLGFTLVK 480

QY 481 PCHEPKCFSSYEYKCKKNNNTGQALPSODEGSKAFVNCDENSRLVSTLNLVTRADG 540  
DB 481 PCHEPKCFSSYEYKCKKNNNTGQALPSODEGSKAFVNCDENSRLVSTLNLVTRADG 540  
QY 541 MYWCGYKOGHFYGETAAVYVAVERKAAGSRVSLAKADAAPEKVDLSGFEIENKAI 600  
DB 541 MYWCGYKOGHFYGETAAVYVAVERKAAGSRVSLAKADAAPEKVDLSGFEIENKAI 600  
QY 601 DPLFAEKAADVADTRDQADGSRASVDSGSSSEOGSSRALVSTVPLGLVLAAGVAVG 660  
DB 601 DPLFAEKAADVADTRDQADGSRASVDSGSSSEOGSSRALVSTVPLGLVLAAGVAVG 660  
QY 661 ARARRKKNDRYSIRSYRDISMSDFENSREFGANDNMKASSITQETSLGKKEEVPATTE 720  
DB 661 ARARRKKNDRYSIRSYRDISMSDFENSREFGANDNMKASSITQETSLGKKEEVPATTE 720  
QY 721 STTETKEPKKAKRSKEEAMAYKDFLLQSSVVAEADQDPOEA 764  
DB 721 STTETKEPKKAKRSKEEAMAYKDFLLQSSVVAEADQDPOEA 764

## RESULT 13

US-60-452-680-13120  
Sequence 13120, Application US/60452680  
GENERAL INFORMATION:  
APPLICANT: CARGILL, Michele  
APPLICANT: GRUPE, Andrew  
TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH  
TITLE OF INVENTION: ALZHEIMER'S DISEASE, METHODS OF DETECTION AND USES THEREOF  
FILE REFERENCE: CL001450  
CURRENT APPLICATION NUMBER: US/60/452, 680  
CURRENT FILING DATE: 2003-03-07  
NUMBER OF SEQ ID NOS: 116213  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 13120  
LENGTH: 764  
TYPE: PRF  
ORGANISM: Homo sapiens  
US-60-452-680-13120

Query Match 86.8%; Score 663; DB 32; Length 764;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 763; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MLFVLTCLAVPAISTKSPIFGPEEVNSVEGNSVITCYPTSVNHRTRKRCROGA 60  
DB 1 MLFVLTCLAVPAISTKSPIFGPEEVNSVEGNSVITCYPTSVNHRTRKRCROGA 60  
QY 61 RGCITLISSEGVSSKYAGRANLTFNPNENGTFVYVNIQLOSDSGRYKCGLGINSRGLS 120  
DB 61 RGCITLISSEGVSSKYAGRANLTFNPNENGTFVYVNIQLOSDSGRYKCGLGINSRGLS 120  
QY 121 FVLSLEVSOGPGLNDTKYTVDLGRTVITNCPFKTENAKRKSILYKQIGLPLVYIDS 180  
DB 121 FVLSLEVSOGPGLNDTKYTVDLGRTVITNCPFKTENAKRKSILYKQIGLPLVYIDS 180  
QY 181 GYVNPYNGRIRLDIOGTGQLFSVYINQLRLSDAGQYLCQAGDSDNSNKKNDLOVLPK 240  
DB 181 GYVNPYNGRIRLDIOGTGQLFSVYINQLRLSDAGQYLCQAGDSDNSNKKNDLOVLPK 240  
QY 241 EPELVYEDLRGSVTFHCALGPEVANYAKFLCROSSGENDVYVNLGKRAPAFEGRIILN 300  
DB 241 EPELVYEDLRGSVTFHCALGPEVANYAKFLCROSSGENDVYVNLGKRAPAFEGRIILN 300  
QY 301 POKKDSFSVYITGLRKEDAGRYLCGAHSDGLOEGSPIOAMOLFVNEESTIPRSPTVYK 360  
DB 301 POKKDSFSVYITGLRKEDAGRYLCGAHSDGLOEGSPIOAMOLFVNEESTIPRSPTVYK 360  
QY 361 GYAGSSVAVLCPYNRKESKSIKYWCLMEGAONGRCPPLVDSGVAQYEGRLSLEBPG 420  
DB 361 GYAGSSVAVLCPYNRKESKSIKYWCLMEGAONGRCPPLVDSGVAQYEGRLSLEBPG 420  
QY 421 NGFTFYILNQLSRDAGFYWCLTNGDTLMRTTVEIKIIEGEPNLKVPNGVTAVLGFTLVK 480



|||||  
Db 421 NGFTVILNQLTSDRAGFYWCLTNGDITLMTTVEIKIIEGEPNLKVPGNVAVLGETLKV 480  
Oy 481 PCHPCKFSSYEKWKCKNNNGCOALPSQDEGPKSAFVNCDENSRLVSLTINLYTRADEG 540  
Db 481 PCHPCKFSSYEKWKCKNNNGCOALPSQDEGPKSAFVNCDENSRLVSLTINLYTRADEG 540  
Oy 541 WYMGVKGHFGYGETAAVYVAVEERKAAGSDVSLAKADAAPEKVLDSGFEIENKA1Q 600  
Db 541 WYMGVKGHFGYGETAAVYVAVEERKAAGSDVSLAKADAAPEKVLDSGFEIENKA1Q 600  
Oy 601 DPLFAEKAADTRDQADGSRASVDSGSSSEOGSSRALVSTLVLPLGLVAVGAVGV 660  
Db 601 DPLFAEKAADTRDQADGSRASVDSGSSSEOGSSRALVSTLVLPLGLVAVGAVGV 660  
Oy 661 ARARRKKNVDRVSTIRSTRD1SMSPFENSREGANDNNGASSITQETSLGKEEFVATTE 720  
Db 661 ARARRKKNVDRVSTIRSTRD1SMSPFENSREGANDNNGASSITQETSLGKEEFVATTE 720  
Oy 721 STTEKPKKAKRSSKEEAEMAYKDFLLQSSVVAEADGPOEA 764  
Db 721 STTEKPKKAKRSSKEEAEMAYKDFLLQSSVVAEADGPOEA 764

RESULT 14  
US-09-949-016-8119

Sequence 8119, Application US/09949016  
GENERAL INFORMATION:  
APPLICANT: VENTER, J. Craig et al.  
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
FILE REFERENCE: C1001307  
CURRENT APPLICATION NUMBER: US/09/949,016  
PRIOR FILING DATE: 2000-04-14  
PRIOR APPLICATION NUMBER: 60/241,755  
PRIOR FILING DATE: 2000-10-20  
PRIOR APPLICATION NUMBER: 60/237,768  
PRIOR FILING DATE: 2000-10-03  
PRIOR APPLICATION NUMBER: 60/231,498  
PRIOR FILING DATE: 2000-09-08  
NUMBER OF SEQ ID NOS: 207012  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 8119  
LENGTH: 816  
TYPE: PRT  
ORGANISM: Human  
US-09-949-016-8119

Query Match 86.8%; Score 663; DB 24; Length 816;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 763; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 1 MLPLVTCLLAVFAPISTKSPFGEPEVNSVEGNSVSTCYPPPTSVMRHRKRYWCRGA 60  
Db 53 MLPLVTCLLAVFAPISTKSPFGEPEVNSVEGNSVSTCYPPPTSVMRHRKRYWCRGA 112  
Oy 61 RGGCTTLISSEGYSSKAGRANLNFENGTFVYVNIQALSDODSGRCKGIGINSRLS 120  
Db 113 RGGCTTLISSEGYSSKAGRANLNFENGTFVYVNIQALSDODSGRCKGIGINSRLS 172  
Oy 121 FDSVLEVSQGPGLNDKTVYTVVDLGRVTYVINCPEKTENAQRKSLYKOIGLYPVLVIDSS 180  
Db 173 FDSVLEVSQGPGLNDKTVYTVVDLGRVTYVINCPEKTENAQRKSLYKOIGLYPVLVIDSS 232  
Oy 181 GYVNPNTYGRIRLDIOGQGLFSVYINQLRLSDAGQYLCQAGDSDSNKKNKADQVILKP 240  
Db 233 GYVNPNTYGRIRLDIOGQGLFSVYINQLRLSDAGQYLCQAGDSDSNKKNKADQVILKP 292  
Oy 241 EPELVYEDLRGCVTFHACALGPEVANVAKFLCROSSGENDVYVNTLGRAPAFEGRIILN 300  
Db 293 EPELVYEDLRGCVTFHACALGPEVANVAKFLCROSSGENDVYVNTLGRAPAFEGRIILN 352  
Oy 301 PODKDGFSFVYITGLRKEDAGRYLCGANSQQLQEGSPIQAWOLFVNEESTIPRSPYVVK 360

|||||  
Db 353 PODKDGFSFVYITGLRKEDAGRYLCGANSQQLQEGSPIQAWOLFVNEESTIPRSPYVVK 412  
Oy 361 GYASSVAVLCPPYRKESKSTIKYWCLEMGANCGCPLLJOSEGVNKAQYEGRLSLLEPG 420  
Db 413 GVAGSVAVLCPPYRKESKSTIKYWCLEMGANCGCPLLJOSEGVNKAQYEGRLSLLEPG 472  
Oy 421 NGFTVILNQLTSDRAGFYWCLTNGDITLMTTVEIKIIEGEPNLKVPGNVAVLGETLKV 480  
Db 473 NGFTVILNQLTSDRAGFYWCLTNGDITLMTTVEIKIIEGEPNLKVPGNVAVLGETLKV 532  
Oy 481 PCHPCKFSSYEKWKCKNNNGCOALPSQDEGPKSAFVNCDENSRLVSLTINLYTRADEG 540  
Db 533 PCHPCKFSSYEKWKCKNNNGCOALPSQDEGPKSAFVNCDENSRLVSLTINLYTRADEG 592  
Oy 541 WYMGVKGHFGYGETAAVYVAVEERKAAGSDVSLAKADAAPEKVLDSGFEIENKA1Q 600  
Db 593 WYMGVKGHFGYGETAAVYVAVEERKAAGSDVSLAKADAAPEKVLDSGFEIENKA1Q 652  
Oy 601 DPLFAEKAADTRDQADGSRASVDSGSSSEOGSSRALVSTLVLPLGLVAVGAVGV 660  
Db 653 DPLFAEKAADTRDQADGSRASVDSGSSSEOGSSRALVSTLVLPLGLVAVGAVGV 712  
Oy 661 ARARRKKNVDRVSTIRSTRD1SMSPFENSREGANDNNGASSITQETSLGKEEFVATTE 720  
Db 713 ARARRKKNVDRVSTIRSTRD1SMSPFENSREGANDNNGASSITQETSLGKEEFVATTE 772  
Oy 721 STTEKPKKAKRSSKEEAEMAYKDFLLQSSVVAEADGPOEA 764  
Db 773 STTEKPKKAKRSSKEEAEMAYKDFLLQSSVVAEADGPOEA 816

## RESULT 15

US-09-950-294-4  
Sequence 4, Application US/09950294

GENERAL INFORMATION:

APPLICANT: Moritson, Sherie L.  
Chintalacharuvu, Kote R.

TITLE OF INVENTION: SECRETORY IMMUNOGLOBULIN PRODUCED  
BY SINGLE CELLS AND METHODS FOR MAKING AND USING  
SAME

NUMBER OF SEQUENCES: 4

CORRESPONDENCE ADDRESS:

ADDRESS: Merchant, Gould, Smith, Edell, Welter & Schmidt  
STREET: 11150 Santa Monica Boulevard, Suite 400  
CITY: Los Angeles

STATE: CA

COUNTRY: USA

ZIP: 90025

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

SOFTWARE: FastSeq for Windows Version 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/950,294

FILING DATE: 10-Sep-2001

CLASSIFICATION: <unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 09/095,385

FILING DATE: <unknown>

ATTORNEY/AGENT INFORMATION:

NAME: Canada, Karen S

REGISTRATION NUMBER: 39,927

REFERENCE/DOCKET NUMBER: 30435,45USU1

TELECOMMUNICATION INFORMATION:

TELEPHONE: 310 445-1140

TELEFAX: 310 445-9031

TELEX: <unknown>

INFORMATION FOR SEQ ID NO: 4:

SEQUENCE CHARACTERISTICS:

LENGTH: 608 amino acids

TYPE: amino acid

STRANDEDNESS: unknown  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 4:  
US-09-950-294-4

Query Match 79.5%; Score 607; DB 25; Length 608;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 607; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 MLFVLTCLAVPAISTKSPIRGPPEVNSVEGNSVSYTCYPTSVNRRTRKYWCROGA 60
   |||||||
Db 1 MLFVLTCLAVPAISTKSPIRGPPEVNSVEGNSVSYTCYPTSVNRRTRKYWCROGA 60
   |||||||
QY 61 RGCITLISSEGYSSKRYAGRANLTNFPENGTFVNI AOLSDDSGRYKCGLGINSRGLS 120
   |||||||
Db 61 RGCITLISSEGYSSKRYAGRANLTNFPENGTFVNI AOLSDDSGRYKCGLGINSRGLS 120
   |||||||
QY 121 FDVSLVSYSGPGLNDTKYTVTDGRTVTINCPFKTENAKRSLYKQIGLYPVLYDSS 180
   |||||||
Db 121 FDVSLVSYSGPGLNDTKYTVTDGRTVTINCPFKTENAKRSLYKQIGLYPVLYDSS 180
   |||||||
QY 181 GYVNPNTGRIRLDIOGTGQLFSVYINOLRLSDAGQYLCQAGDDSNKKNADLOYLKP 240
   |||||||
Db 181 GYVNPNTGRIRLDIOGTGQLFSVYINOLRLSDAGQYLCQAGDDSNKKNADLOYLKP 240
   |||||||
QY 241 EPELVYEDLRGSYTFHCALGPEVANNAKFLCRQSSGENCDVYNTLGKRAPEGRILLN 300
   |||||||
Db 241 EPELVYEDLRGSYTFHCALGPEVANNAKFLCRQSSGENCDVYNTLGKRAPEGRILLN 300
   |||||||
QY 301 PODKDSFSYVTGLRKEDAGRYLCGAHSDGQLQEGSPIQAMOLFVNEESTIRSPYVVK 360
   |||||||
Db 301 PODKDSFSYVTGLRKEDAGRYLCGAHSDGQLQEGSPIQAMOLFVNEESTIRSPYVVK 360
   |||||||
QY 361 GYAGSSYAVYLCPIYNRKESKIKYWCLEGAONGRCPLVDSSEGWKAQYEGRLSLEEPG 420
   |||||||
Db 361 GYAGSSYAVYLCPIYNRKESKIKYWCLEGAONGRCPLVDSSEGWKAQYEGRLSLEEPG 420
   |||||||
QY 421 NGFTYVILNQLTSRDAGFYWCLTNGDTLMRTVEIKIIEGEPNLKYPGNTAVLGETLKV 480
   |||||||
Db 421 NGFTYVILNQLTSRDAGFYWCLTNGDTLMRTVEIKIIEGEPNLKYPGNTAVLGETLKV 480
   |||||||
QY 481 PCHFPECKFSSEYKWKCMNNTGQALPSODEGSKAFVNCDENSRLVSLTNLVTRADEG 540
   |||||||
Db 481 PCHFPECKFSSEYKWKCMNNTGQALPSODEGSKAFVNCDENSRLVSLTNLVTRADEG 540
   |||||||
QY 541 WYMGVYKQGHFYGETAAVYAVEERKAGSRDYSIAKADAAPDEKYLDSGFREIENKAIQ 600
   |||||||
Db 541 WYMGVYKQGHFYGETAAVYAVEERKAGSRDYSIAKADAAPDEKYLDSGFREIENKAIQ 600
   |||||||
QY 601 DPLRLFAE 607
   |||||||
Db 601 DPLRLFAE 607
   |||||||
```

Search completed: September 9, 2003, 23:28:01  
Job time : 398 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 9, 2003, 23:17:53 ; Search time 15 Seconds

(without alignments)  
486.619 Million cell updates/sec

Title: US-09-818-247-1

Sequence: 1 MLFVLTCLAVPAISTKS.....DFLQSTVAEAADGPOEA 764

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 47124 seqs, 9554045 residues

Word size : 0

Total number of hits satisfying chosen parameters: 47124

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

Pending\_Patents\_AA\_New:\*  
1: /cgn2\_6/ptodata/1/paa/PCT\_NEW\_COMB.pep:\*  
2: /cgn2\_6/ptodata/1/paa/US06\_NEW\_COMB.pep:\*  
3: /cgn2\_6/ptodata/1/paa/US07\_NEW\_COMB.pep:\*  
4: /cgn2\_6/ptodata/1/paa/US08\_NEW\_COMB.pep:\*  
5: /cgn2\_6/ptodata/1/paa/US09\_NEW\_COMB.pep:\*  
6: /cgn2\_6/ptodata/1/paa/US10\_NEW\_COMB.pep:\*  
7: /cgn2\_6/ptodata/1/paa/US60\_NEW\_COMB.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	16	2.1	16	US-09-005-318F-45	Sequence 45, Appl
2	16	2.1	16	US-10-062-467B-45	Sequence 45, Appl
3	8	1.0	762	PCT-US02-29560A-285	Sequence 285, App
4	8	1.0	802	PCT-US02-29560A-284	Sequence 284, App
5	8	1.0	1308	US-60-487-610-1446	Sequence 1446, Ap
6	8	1.0	1324	US-60-487-610-1445	Sequence 1445, Ap
7	8	1.0	1379	US-60-487-610-1444	Sequence 1444, Ap
8	7	0.9	21	US-60-490-419-86	Sequence 86, Appl
9	7	0.9	21	US-60-490-149-86	Sequence 86, Appl
10	7	0.9	209	US-60-495-114-2042	Sequence 2042, Ap
11	7	0.9	216	PCT-US03-25276-32	Sequence 32, Appl
12	7	0.9	287	PCT-US02-18638A-176	Sequence 176, Appl
13	7	0.9	300	PCT-US02-18638A-174	Sequence 174, Appl
14	7	0.9	300	PCT-US03-10783-2	Sequence 2, Appl
15	7	0.9	300	US-60-487-610-2568	Sequence 2568, App
16	7	0.9	314	PCT-US02-18638A-172	Sequence 172, Appl
17	7	0.9	314	US-09-981-845-1	Sequence 1, Appl
18	7	0.9	314	US-60-487-610-2567	Sequence 2567, Ap
19	7	0.9	336	US-09-897-516A-7391	Sequence 7391, Ap
20	7	0.9	404	PCT-US03-21559-7	Sequence 7, Appl
21	7	0.9	479	US-60-478-196-3256	Sequence 3256, App
22	7	0.9	603	PCT-US02-18638A-186	Sequence 186, App
23	7	0.9	603	US-60-490-890-1440	Sequence 1440, App
24	7	0.9	630	PCT-US02-29560A-364	Sequence 364, App
25	7	0.9	653	US-60-485-450-1067	Sequence 1067, App
26	7	0.9	687	US-60-485-450-1066	Sequence 1066, App

27	7	0.9	851	US-09-876-773-2	Sequence 2, Appl
28	7	0.9	851	US-60-487-610-2240	Sequence 2240, Ap
29	7	0.9	851	US-60-485-450-1426	Sequence 1426, Ap
30	7	0.9	866	PCT-US02-37660-6	Sequence 6, Appl
31	7	0.9	898	US-60-478-196-3332	Sequence 3332, Ap
32	7	0.9	1331	PCT-US03-22467-2	Sequence 2, Appl
33	7	0.9	1333	PCT-US03-22467-12	Sequence 12, Appl
34	7	0.9	1798	US-60-487-610-2811	Sequence 2811, Ap
35	7	0.9	1798	US-60-487-610-2812	Sequence 2812, Ap
36	7	0.9	1954	PCT-US03-19027-2	Sequence 2, Appl
37	7	0.9	3714	US-60-487-610-3753	Sequence 2753, Ap
38	6	0.8	6	US-09-005-318F-41	Sequence 41, Appl
39	6	0.8	6	US-10-062-467B-41	Sequence 41, Appl
40	6	0.8	20	PCT-US03-11766-5	Sequence 5, Appl
41	6	0.8	34	PCT-US03-18657-3	Sequence 3, Appl
42	6	0.8	36	PCT-US03-18657-2	Sequence 2, Appl
43	6	0.8	62	US-10-644-807-251	Sequence 251, App
44	6	0.8	62	US-10-644-807-345	Sequence 345, App
45	6	0.8	82	PCT-US03-25418-27	Sequence 27, Appl

## ALIGNMENTS

```
RESULT 1
US-09-005-318F-45
; Sequence 45, Application US/09005318F
; GENERAL INFORMATION:
; APPLICANT: HEIN, MICH B.
; APPLICANT: HINAT, ANDREW C.
; TITLE OF INVENTION: NOVEL EPITHELIAL TISSUE TARGETING AGENT
; FILE REFERENCE: EPI3004B
; CURRENT FILING DATE: 1998-01-09
; PRIOR APPLICATION NUMBER: US/09/005,318F
; PRIOR FILING DATE: 1997-01-10
; PRIOR APPLICATION NUMBER: 09/005,167
; PRIOR FILING DATE: 1998-01-09
; NUMBER OF SEQ ID NOS: 113
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 45
; LENGTH: 16
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-005-318F-45

Query Match      2.1% Score 16; DB 5; Length 16;
Best Local Similarity 100.0%; Pred. No. 5; Gaps 0;
Matches 16; Conservative 0; Mismatches 0; Indels 0;

QY      598 AIDPRLFAEKAVAD 613
DB      1 AIDPRLFAEKAVAD 16

RESULT 2
US-10-062-467B-45
; Sequence 45, Application US/10062467B
; GENERAL INFORMATION:
; APPLICANT: HEIN, MICH B.
; APPLICANT: HINAT, ANDREW C.
; TITLE OF INVENTION: J CHAIN POLYPEPTIDE TARGETING MOLECULE LINKED TO AN IMAGING
; FILE REFERENCE: EPI3003C
; CURRENT FILING DATE: 2002-02-05
; PRIOR APPLICATION NUMBER: 08/782,480
; PRIOR FILING DATE: 1997-01-10
; PRIOR APPLICATION NUMBER: 09/005,167
; PRIOR FILING DATE: 1998-01-09
; NUMBER OF SEQ ID NOS: 93
; SOFTWARE: PatentIn Ver. 2.1
```

SEQ ID NO 45  
 LENGTH: 16  
 TYPE: PRT  
 ORGANISM: Homo sapiens  
 US-10-062-467B-45

Query Match 2.18; Score 16; DB 6; Length 16;  
 Best Local Similarity 100.0%; Pred. No. 5.6e-10;  
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 598 AIDDPRLFAEKAAYD 613  
 Db 1 AIDDPRLFAEKAAYD 16

RESULT 3  
 PCT-US02-29560A-285

Sequence 285, Application PC/TUS0229560A  
 GENERAL INFORMATION:

APPLICANT: Afar, Daniel  
 APPLICANT: Aziz, Natasha  
 APPLICANT: Gish, Kurt C.  
 APPLICANT: Hevezl, Peter A.  
 APPLICANT: Mack, David H.  
 APPLICANT: Wilson, Keith E.

APPLICANT: Zlotnik, Albert

TITLE OF INVENTION: Methods of Diagnosis of Cancer, Compositions and  
 TITLE OF INVENTION: Methods of Screening for Modulators of Cancer

FILE REFERENCE: 018501-002710PC

CURRENT APPLICATION NUMBER: PCT/US02/29560A

CURRENT FILING DATE: 2002-09-17

PRIOR APPLICATION NUMBER: US 60/323,469

PRIOR FILING DATE: 2001-09-17

NUMBER OF SEQ ID NOS: 412

SOFTWARE: FastSeq for Windows Version 3.0

SEQ ID NO 285

LENGTH: 762

TYPE: PRT

ORGANISM: Homo sapiens

PCT-US02-29560A-285

Query Match 1.0%; Score 8; DB 1; Length 762;  
 Best Local Similarity 100.0%; Pred. No. 4.1;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 318 EDAGRYLC 325  
 Db 94 EDAGRYLC 101

RESULT 4  
 PCT-US02-29560A-284

Sequence 284, Application PC/TUS0229560A  
 GENERAL INFORMATION:

APPLICANT: Afar, Daniel  
 APPLICANT: Aziz, Natasha  
 APPLICANT: Gish, Kurt C.  
 APPLICANT: Hevezl, Peter A.  
 APPLICANT: Mack, David H.  
 APPLICANT: Wilson, Keith E.  
 APPLICANT: Zlotnik, Albert

TITLE OF INVENTION: Methods of Diagnosis of Cancer, Compositions and  
 TITLE OF INVENTION: Methods of Screening for Modulators of Cancer

FILE REFERENCE: 018501-002710PC

CURRENT APPLICATION NUMBER: PCT/US02/29560A

CURRENT FILING DATE: 2002-09-17

PRIOR APPLICATION NUMBER: US 60/323,469

PRIOR FILING DATE: 2001-09-17

NUMBER OF SEQ ID NOS: 412

SOFTWARE: FastSeq for Windows Version 3.0

SEQ ID NO 284

LENGTH: 802  
 TYPE: PRT  
 ORGANISM: Homo sapiens  
 PCT-US02-29560A-284

Query Match 1.0%; Score 8; DB 1; Length 802;  
 Best Local Similarity 100.0%; Pred. No. 4.4;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 318 EDAGRYLC 325  
 Db 94 EDAGRYLC 101

RESULT 5  
 US-60-487-610-1446

Sequence 1446, Application US/60487610  
 GENERAL INFORMATION:

APPLICANT: CARGILL, Michele  
 APPLICANT: HUANG, Hongjin  
 TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH  
 TITLE OF INVENTION: LIVER FIBROSIS IN HEPATITIS C VIRUS-INFECTED SUBJECTS,  
 TITLE OF INVENTION: METHODS OF DETECTION AND USES THEREOF

FILE REFERENCE: CL001469

CURRENT APPLICATION NUMBER: US/60/487,610

CURRENT FILING DATE: 2003-07-17

NUMBER OF SEQ ID NOS: 97101

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 1446

LENGTH: 1308

TYPE: PRT

ORGANISM: Homo sapiens

US-60-487-610-1446

Query Match 1.0%; Score 8; DB 7; Length 1308;  
 Best Local Similarity 100.0%; Pred. No. 6.9;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 318 EDAGRYLC 325  
 Db 751 EDAGRYLC 758

RESULT 6  
 US-60-487-610-1445

Sequence 1445, Application US/60487610  
 GENERAL INFORMATION:  
 APPLICANT: CARGILL, Michele  
 APPLICANT: HUANG, Hongjin  
 TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH  
 TITLE OF INVENTION: LIVER FIBROSIS IN HEPATITIS C VIRUS-INFECTED SUBJECTS,  
 TITLE OF INVENTION: METHODS OF DETECTION AND USES THEREOF

FILE REFERENCE: CL001469

CURRENT APPLICATION NUMBER: US/60/487,610

CURRENT FILING DATE: 2003-07-17

NUMBER OF SEQ ID NOS: 97101

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 1445

LENGTH: 1324

TYPE: PRT

ORGANISM: Homo sapiens

US-60-487-610-1445

Query Match 1.0%; Score 8; DB 7; Length 1324;  
 Best Local Similarity 100.0%; Pred. No. 7;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 318 EDAGRYLC 325  
 Db 757 EDAGRYLC 764

RESULT 7

```
US-60-487-610-1444
; Sequence 1444, Application US/60487610
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; APPLICANT: HUANG, Hongjin
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: LIVER FIBROSIS IN HEPATITIS C VIRUS-INFECTED SUBJECTS,
; FILE REFERENCE: CL001469
; CURRENT APPLICATION NUMBER: US/60/487,610
; CURRENT FILING DATE: 2003-07-17
; NUMBER OF SEQ ID NOS: 97101
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1444
; LENGTH: 1379
; TYPE: PRF
; ORGANISM: Homo sapiens
US-60-487-610-1444

Query Match
Best Local Similarity 1.0%; Score 8; DB 7; Length 1379;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 318 EDAGRYLC 325
DB 757 EDAGRYLC 764

RESULT 8
US-60-490-419-86
; Sequence 86, Application US/60490419
; GENERAL INFORMATION:
; APPLICANT: Fred E. Regneir
; APPLICANT: Ziang Zhang
; TITLE OF INVENTION: Digital Chromatography
; FILE REFERENCE: 302840.3000-100
; CURRENT APPLICATION NUMBER: US/60/490,419
; CURRENT FILING DATE: 2003-07-25
; NUMBER OF SEQ ID NOS: 277
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 86
; LENGTH: 21
; TYPE: PRF
; ORGANISM: Homo sapiens
US-60-490-419-86

Query Match
Best Local Similarity 0.9%; Score 7; DB 7; Length 21;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 705 QETSIGG 711
DB 8 QETSIGG 14

RESULT 9
US-60-490-149-86
; Sequence 86, Application US/60490149
; GENERAL INFORMATION:
; APPLICANT: Fred E. Regneir
; APPLICANT: Jiri Adamec
; APPLICANT: Ziang Zhang
; TITLE OF INVENTION: Digital Chromatography
; FILE REFERENCE: 302840.3000-100
; CURRENT APPLICATION NUMBER: US/60/490,149
; CURRENT FILING DATE: 2003-07-25
; NUMBER OF SEQ ID NOS: 277
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 86
; LENGTH: 21
; TYPE: PRF
; ORGANISM: Homo sapiens

US-60-490-149-86
Query Match
Best Local Similarity 0.9%; Score 7; DB 7; Length 21;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 705 QETSIGG 711
DB 8 QETSIGG 14

RESULT 10
US-60-495-114-2042
; Sequence 2042, Application US/60495114
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; APPLICANT: CARGILL, Michele
; TITLE OF INVENTION: POLYMORPHISMS IN NUCLEIC ACID MOLECULES
; TITLE OF INVENTION: ENCODING HUMAN PROTEASE PROTEINS, METHODS OF DETECTION AN
; FILE REFERENCE: CL001480
; CURRENT APPLICATION NUMBER: US/60/495,114
; CURRENT FILING DATE: 2003-08-15
; NUMBER OF SEQ ID NOS: 91238
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2042
; LENGTH: 209
; TYPE: PRF
; ORGANISM: Homo sapiens
US-60-495-114-2042

Query Match
Best Local Similarity 0.9%; Score 7; DB 7; Length 209;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 319 DAGRYLC 325
DB 143 DAGRYLC 149

RESULT 11
PCT-US03-25276-32
; Sequence 32, Application PCT/US0325276
; GENERAL INFORMATION:
; APPLICANT: Norris, Susan R
; APPLICANT: Lincoln, Kim
; APPLICANT: Abad, Mark Scott
; APPLICANT: Eilers, Robert
; APPLICANT: Hartsuyker, Karen Kindle
; APPLICANT: Hirschberg, Joseph
; APPLICANT: Karunanandaa, Balasubjini
; APPLICANT: Moshiri, Farhad
; APPLICANT: Stein, Joshua C.
; APPLICANT: Valentin, Henry E.
; APPLICANT: Venkatesh, Tyamagondlu V.
; TITLE OF INVENTION: Tocopherol biosynthesis related genes and used thereof
; FILE REFERENCE: Ren-01-125
; CURRENT APPLICATION NUMBER: PCT/US03/25276
; CURRENT FILING DATE: 2003-08-05
; PRIOR APPLICATION NUMBER: US 60/400,689
; PRIOR FILING DATE: 2002-08-05
; NUMBER OF SEQ ID NOS: 79
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 32
; LENGTH: 216
; TYPE: PRF
; ORGANISM: Synechococcus sp.-WH-8102
PCT-US03-25276-32

Query Match
Best Local Similarity 0.9%; Score 7; DB 1; Length 216;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 639 ALVSTLV 645
```

Db 157 ALVSTLV 163

## RESULT 12

PCT-US02-18638A-176  
; Sequence 176, Application PC/TUS0218638A  
; GENERAL INFORMATION:  
; APPLICANT: Millennium Pharmaceuticals, Inc. et al.  
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR  
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY  
; FILE REFERENCE: MRI-035PC  
; CURRENT APPLICATION NUMBER: PCT/US02/18638A  
; CURRENT FILING DATE: 2002-06-12  
; PRIOR APPLICATION NUMBER: US 60/298,159  
; PRIOR FILING DATE: 2001-06-13  
; PRIOR APPLICATION NUMBER: US 60/298,155  
; PRIOR FILING DATE: 2001-06-13  
; PRIOR APPLICATION NUMBER: US 60/335,936  
; PRIOR FILING DATE: 2001-11-14  
; NUMBER OF SEQ ID NOS: 238  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 176  
; LENGTH: 287  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
PCT-US02-18638A-176

Query Match 0.9%; Score 7; DB 1; Length 287;  
Best Local Similarity 100.0%; Pred. No. 18;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 626 DSGSSEE 632  
Db 23 DSGSSEE 29

## RESULT 13

PCT-US02-18638A-174  
; Sequence 174, Application PC/TUS0218638A  
; GENERAL INFORMATION:  
; APPLICANT: Millennium Pharmaceuticals, Inc. et al.  
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR  
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY  
; FILE REFERENCE: MRI-035PC  
; CURRENT APPLICATION NUMBER: PCT/US02/18638A  
; CURRENT FILING DATE: 2002-06-12  
; PRIOR APPLICATION NUMBER: US 60/298,159  
; PRIOR FILING DATE: 2001-06-13  
; PRIOR APPLICATION NUMBER: US 60/298,155  
; PRIOR FILING DATE: 2001-06-13  
; PRIOR APPLICATION NUMBER: US 60/335,936  
; PRIOR FILING DATE: 2001-11-14  
; NUMBER OF SEQ ID NOS: 238  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 174  
; LENGTH: 300  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
PCT-US02-18638A-174

Query Match 0.9%; Score 7; DB 1; Length 300;  
Best Local Similarity 100.0%; Pred. No. 19;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 626 DSGSSEE 632  
Db 23 DSGSSEE 29

## RESULT 14

PCT-US03-10783-2  
; Sequence 2, Application PC/TUS0310783  
; GENERAL INFORMATION:  
; APPLICANT: Wang, Xin Wei  
; APPLICANT: Ye, Qing-Hai  
; APPLICANT: Kim, Jin Woo  
; APPLICANT: The Government of the United States of America  
; APPLICANT: as represented by the Secretary of the  
; TITLE OF INVENTION: Methods of Diagnosing Potential for Metastasis or  
; TITLE OF INVENTION: Developing Hepatocellular Carcinoma and of Identifying  
; FILE REFERENCE: 015280-463100PC  
; CURRENT APPLICATION NUMBER: PCT/US03/10783  
; CURRENT FILING DATE: 2000-04-04  
; PRIOR APPLICATION NUMBER: US 60/370,895  
; PRIOR FILING DATE: 2002-04-05  
; NUMBER OF SEQ ID NOS: 12  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 2  
; LENGTH: 300  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; OTHER INFORMATION: osteopontin (OPN)  
PCT-US03-10783-2

Query Match 0.9%; Score 7; DB 1; Length 300;  
Best Local Similarity 100.0%; Pred. No. 19;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 626 DSGSSEE 632  
Db 23 DSGSSEE 29

## RESULT 15

US-60-487-610-2568  
; Sequence 2568, Application US/60487610  
; GENERAL INFORMATION:  
; APPLICANT: CARGILL, Michele  
; APPLICANT: HUANG, Hongjin  
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH  
; TITLE OF INVENTION: LIVER FIBROSIS IN HEPATITIS C VIRUS-INFECTED SUBJECTS,  
; TITLE OF INVENTION: METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001469  
; CURRENT APPLICATION NUMBER: US/60/487,610  
; CURRENT FILING DATE: 2003-07-17  
; NUMBER OF SEQ ID NOS: 97101  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 2568  
; LENGTH: 300  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-60-487-610-2568

Query Match 0.9%; Score 7; DB 7; Length 300;  
Best Local Similarity 100.0%; Pred. No. 19;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 626 DSGSSEE 632  
Db 23 DSGSSEE 29

Search completed: September 9, 2003, 23:28:24  
Job time : 19 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 9, 2003, 23:14:22; Search time 45 Seconds

(without alignments)  
1632.731 Million cell updates/sec

Title: US-09-818-247-1

Perfect score: 764  
Sequence: 1 MLEVLCTCLAVFAIPATSKS.....DFLLQSTVAEAADQDQEA 764

Scoring table: OLIGO  
Gapop 60.0, Gapext 60.0

Searched: 283308 seqs, 96168682 residues

Word size: 0

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database:

PIR\_76:\*  
1: PIR1:\*  
2: PIR2:\*  
3: PIR3:\*  
4: PIR4:\*

Pred. NO. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	764	100.0	764	1	ORHUGS secretory componen
2	21	2.7	757	1	S48841 secretory componen
3	21	2.7	757	1	I45956 polymeric immunogl
4	17	2.2	769	1	ORRTGS secretory componen
5	13	1.7	773	1	ORRBG secretory componen
6	10	1.3	308	2	B84311 hypothetical prote
7	8	1.0	48	2	T35253 small hypothetical
8	8	1.0	56	2	E64402 hypothetical prote
9	8	1.0	197	2	F71248 probable proteasom
10	8	1.0	251	2	E90206 hypothetical prote
11	8	1.0	287	2	S55662 legument protein 6
12	8	1.0	287	2	A82318 conserved hypotet
13	8	1.0	308	2	T45133 restriction endonu
14	8	1.0	400	2	D64462 hypothetical prote
15	8	1.0	528	2	T21834 hypothetical prote
16	8	1.0	556	2	E75049 phenylalanine-tRNA
17	8	1.0	694	2	A69768 transcriptipion anti
18	8	1.0	799	2	S18209 fibroblast growth
19	8	1.0	802	1	TVHUF4 fibroblast growth
20	8	1.0	1298	2	A48999 protein-tyrosine k
21	8	1.0	1363	2	I58375 protein-tyrosine k
22	8	1.0	1715	2	C40228 neurexin II-alpha
23	7	0.9	36	2	A40723 poly-ig receptor -
24	7	0.9	87	2	S43906 hypothetical prote
25	7	0.9	89	2	A84301 hypothetical prote
26	7	0.9	93	2	AB2212 hypothetical prote
27	7	0.9	94	2	S48421 probable membrane
28	7	0.9	103	2	A32167 homeotic protein H
29	7	0.9	110	2	T31045 hypothetical prote

30	7	0.9	111	2	G85630 hypothetical prote
31	7	0.9	113	2	F84671 hypothetical prote
32	7	0.9	116	2	D90007 conserved hypotet
33	7	0.9	118	2	S59930 hypothetical prote
34	7	0.9	119	2	T34624 NADH2 dehydrogenas
35	7	0.9	120	2	G70700 hypothetical prote
36	7	0.9	125	2	T05445 hypothetical prote
37	7	0.9	126	2	D95410 hypothetical prote
38	7	0.9	127	2	E86158 60S ribosomal prot
39	7	0.9	128	2	H70435 hypothetical prote
40	7	0.9	128	2	H97140 probable membrane
41	7	0.9	135	2	F87515 hypothetical prote
42	7	0.9	138	2	A81264 probable periplasm
43	7	0.9	142	2	A59098 hypothetical prote
44	7	0.9	143	1	KEMS epsilon-casein pre
45	7	0.9	146	2	E89842 hypothetical prote

#### ALIGNMENTS

##### RESULT 1

secretory component precursor [validated] - human

N:Alternate names: poly-ig receptor; polymeric immunoglobulin receptor

N:Contains: free secretory component; transmembrane secretory component

C:Species: Homo sapiens (man)

C>Date: 28-Aug-1985 #sequence-revision 23-Aug-1996 #text-change 08-Dec-2000

C:Accession: A46537; A55284; I38115; A32263; S38978; S13453; A02112

R:Krajci, P.; Kvale, D.; Tasken, K.; Brandtzaeg, P.

Eur. J. Immunol. 22, 2309-2315, 1992

A:Title: Molecular cloning and exon-intron mapping of the gene encoding human tra

A:Reference number: A46537; MUID:92387236; PMID:135431

A:Accession: A46537

A>Status: not compared with conceptual translation

A:Molecule type: DNA

A:Residues: 1-764 <KRA>

A:Cross-references: GB:S43449; NID:9255097; PIDN:AA23176.1; PID:9255098

A:Experimental source: leukocytes

A>Note: sequence extracted from NCBI backbone (NCBIP:113253)

R:Krajci, P.; Grzeschik, K.H.; Geurts van Kessel, A.H.; Olausen, B.; Brandtzaeg, J

Hum. Genet. 87, 642-648, 1991

A:Title: The human transmembrane secretory component (poly-ig receptor): molecula

A:Reference number: A55284; MUID:92039621; PMID:1682231

A:Accession: A55284

A:Molecule type: mRNA

A:Residues: 1-764 <KRA>

A:Cross-references: GB:S62403; NID:9238235; PIDN:AA20203.1; PID:9238236

A:Experimental source: colonic adenocarcinoma cell line

A>Note: sequence extracted from NCBI backbone (NCBIP:62403; NCBIP:62408)

R:Piskurich, J.F.; France, J.A.; Tamer, C.M.; Willmer, C.A.; Kaetzl, C.S.; Kaetzl

Mol. Immunol. 30, 413-421, 1993

A:Title: Interferon-gamma induces polymeric immunoglobulin receptor mRNA in human

A:Reference number: I38115; MUID:93205018; PMID:8455639

A:Accession: I38115

A:Molecule type: mRNA

A:Residues: 1-764 <RES>

A:Cross-references: EMBL:X73079; NID:9456345; PIDN:CAA51532.1; PID:9456346

A>Note: submitted to the EMBL/GenBank/DBJ databases by J.F. Piskurich, February

R:Krajci, P.; Solberg, R.; Sandberg, M.; Oyen, O.; Johnsen, T.; Brandtzaeg, P.

Biochem. Biophys. Res. Commun. 158, 783-789, 1989

A:Title: Molecular cloning of the human transmembrane secretory component (poly-ig

A:Reference number: A32263; MUID:89149795; PMID:2920039

A:Accession: A32263

A:Molecule type: mRNA

A:Residues: 72-764 <KR3>

A:Cross-references: GB:M24559; NID:9514365; PIDN:AA36102.1; PID:9514366

R:Fallgren-Gebauer, E.; Gebauer, W.; Bastian, A.; Kratzin, H.D.; Elffert, H.; Zhi

Biol. Chem. Hoppe-Seyler 374, 1023-1028, 1993

A:Title: The covalent linkage of secretory component to IgA. Structure of sIgA.

A:Reference number: S38978; MUID:94121784; PMID:8292260

A:Accession: S38978

A:Molecule type: protein

A:Residues: 478-488:517-526:543-545 <PAL>  
A:Note: disulfide bonds for unbound and Iga-bound forms  
R:Elfert, H.; Quentlin, E.; Wiedehof, M.; Hillemeier, S.; Decker, J.; Weber, M.; Hilsch  
Biochem. Hoppe-Seyler 372, 119-128, 1991  
A:Title: Determination of the molecular structure of the human free secretory component.  
A:Reference number: S13453; MUID:91315750; PMID:1859628  
A:Accession: S13453  
A:Molecule type: protein  
A:Residues: 19-157, 'Q', 137-157, 'D', 159-207, 'DE', 210-228, 230-233, 'N', 235-240, 'Q', 242-261,  
R:Elfert, H.; Quentlin, E.; Decker, J.; Hillemeier, S.; Hufschmidt, M.; Klingmüller, D.;  
Hoppe-Seyler's Z. Physiol. Chem. 365, 1489-1495, 1984  
A:Title: The primary structure of the human free secretory component and the arrangement  
A:Reference number: A02112; MUID:65128981; PMID:6526384  
A:Accession: A02112  
A:Molecule type: protein  
A:Residues: 19-157, 'D', 159-207, 'DE', 210-228, 230-233, 'N', 235-240, 'Q', 242-261,  
A:Note: paper in German with English abstract  
C:Comment: As a 100k transmembrane receptor for polymeric immunoglobulins, secretory com-  
plicated, forms interchain disulfide bonds, undergoes proteolysis and transcytosis. Free s  
C:Genetics:  
A:Gene: GDB:PIGR  
A:Cross-references: GDB:120290; OMIM:173880  
A:Map position: 1q31-1q41  
A:Introns: 15/1: 130/1: 349/1: 460/1: 569/1: 629/2: 670/1: 714/1: 733/3  
A:Note: the first intron occurs before the initiator codon  
C:Complex: monomeric as a transmembrane receptor or free in mucosal secretions; heterod-  
amers; hetero-22-mer composed of one chain of secretory component, one chain of immunog-  
C:Superfamily: secretory component; immunoglobulin homology  
C:Keywords: duplication; glycoprotein; immunoglobulin receptor; phosphoprotein; transcyt  
F:1-18/Domain: signal sequence #status predicted <Sig>  
F:19-764/Product: transmembrane secretory component #status predicted <MATM>  
F:19-577/Product: free secretory component #status experimental <MATF>  
F:33-112/Domain: immunoglobulin homology <IM1>  
F:145-222/Domain: immunoglobulin homology <IM2>  
F:250-327/Domain: immunoglobulin homology <IM3>  
F:364-443/Domain: immunoglobulin homology <IM4>  
F:475-546/Domain: immunoglobulin homology <IM5>  
F:639-661/Domain: transmembrane #status predicted <TMN>  
F:662-764/Domain: intracellular #status predicted <IN>  
F:40-110, 135, 186, 421, 469, 499/Binding site: carbohydrate (asn) (covalent) #status experim  
F:486-520/Disulfide bonds: (in Ig-unbound form) #status experimental  
F:486-520/Disulfide bonds: interchain (to Iga alpha-1 chain-192) #status experimental  
F:520/Binding site: cysteine (Cys) (covalent) (in Ig-bound form) #status experimental  
F:577-578/Cleavage site: Lys-Ala (unidentified protease) #status experimental  
F:673/Binding site: phosphate (Ser) (covalent) #status predicted

Query Match 100.0%: Score 764: DB 1: Length 764:  
Best Local Similarity 100.0%: Pred. No. 0:  
Matches 764: Conservative 0: Mismatches 0: Indels 0: Gaps 0:

QY 1 MLFVLTCLAVEPAISTKSPIFGPEEVSNEGSNSVITCYPTSVNRHTRKYWCROGA 60  
Db 1 MLFVLTCLAVEPAISTKSPIFGPEEVSNEGSNSVITCYPTSVNRHTRKYWCROGA 60  
QY 61 RGCITLISSEGVSSKRYARANTLNPENGTFVYVIAQLSODDSGRYKGLGINSGLS 120  
Db 61 RGCITLISSEGVSSKRYARANTLNPENGTFVYVIAQLSODDSGRYKGLGINSGLS 120  
QY 121 FVDSLEVSQGPLNDKRYTVYDGRVTINCPKTEKNAKRSLYQICLTPVLYDSS 180  
Db 121 FVDSLEVSQGPLNDKRYTVYDGRVTINCPKTEKNAKRSLYQICLTPVLYDSS 180  
QY 181 GYVNPNTGRIRLDIOGTGQLFSLVYINQLRLSDAGQYLCQAGDDSNKRNKMDLOYLKP 240  
Db 181 GYVNPNTGRIRLDIOGTGQLFSLVYINQLRLSDAGQYLCQAGDDSNKRNKMDLOYLKP 240  
QY 241 EPELYEDLGSVTFHCLGPEVANAQKFLCROSSGNCVYVNTLGRAPAEGRILN 300  
Db 241 EPELYEDLGSVTFHCLGPEVANAQKFLCROSSGNCVYVNTLGRAPAEGRILN 300  
QY 301 PODKDSFSVYITGLRKEDAGRYLCGAHSDGQLEQSPIDAMOLFVNEESTIRSPVYK 360  
Db 301 PODKDSFSVYITGLRKEDAGRYLCGAHSDGQLEQSPIDAMOLFVNEESTIRSPVYK 360

Db 301 PODKDSFSVYITGLRKEDAGRYLCGAHSDGQLEQSPIDAMOLFVNEESTIRSPVYK 360  
QY 361 GVAGSSVAVLCPPYRNRESKIKYWCMEGANOGRCLYDSEGVNAQYEGRLSLEBPG 420  
Db 361 GVAGSSVAVLCPPYRNRESKIKYWCMEGANOGRCLYDSEGVNAQYEGRLSLEBPG 420  
QY 421 NGFTFYILNDLSDRDAGFYWCLTNGTLMRTTYEITIGEPRLKYPGVNTAVLGETLV 480  
Db 421 NGFTFYILNDLSDRDAGFYWCLTNGTLMRTTYEITIGEPRLKYPGVNTAVLGETLV 480  
QY 481 PCGFPEKFSSEYKWKMNNTGQALPDSODEGSKAFVNCDENSRLVSLTLNLYTADSG 540  
Db 481 PCGFPEKFSSEYKWKMNNTGQALPDSODEGSKAFVNCDENSRLVSLTLNLYTADSG 540  
QY 541 WYMGVYKQGHFYGETAAYVAVEERRAAGSRDYSKADAPDEKVLDSGFREIENKA10 600  
Db 541 WYMGVYKQGHFYGETAAYVAVEERRAAGSRDYSKADAPDEKVLDSGFREIENKA10 600  
QY 601 DPLFLPEERAVATROADGSRASVSGSSEEDGGSSRLVSLTVPLGLAVGANVGY 660  
Db 601 DPLFLPEERAVATROADGSRASVSGSSEEDGGSSRLVSLTVPLGLAVGANVGY 660  
QY 661 ARARHRRKNDVRSIRSYRTDISMSDEPNSEFGANDMGA51TQETSLGKBEFVATTE 720  
Db 661 ARARHRRKNDVRSIRSYRTDISMSDEPNSEFGANDMGA51TQETSLGKBEFVATTE 720  
QY 721 STTETKEPKKAKSSKEEEMAYKDFLLQSTVAAEAODGPOEA 764  
Db 721 STTETKEPKKAKSSKEEEMAYKDFLLQSTVAAEAODGPOEA 764

RESULT 2  
S48841  
secretory component precursor - bovine  
N:Alternate names: poly-Ig receptor; polymeric immunoglobulin receptor  
C:Contains: free secretory component; transmembrane secretory component  
C:Species: Bos primigenius taurus (cattle)  
C:Date: 10-Sep-1999 #sequence, revision 10-Sep-1999 #text, change 10-Sep-1999  
R:Vermeer, H.; Warmerdam, G.W.; de Boer, H.A.; Verbeet, M.P.H.  
submitted to the EMBL Data Library, September 1994  
A:Description: The cloning, tissue specific expression and interspecies sequence c  
A:Reference number: S48841  
A:Accession: S48841  
A:Molecule type: mRNA  
A:Residues: 1-757 <VER>  
A:Cross-references: EMBL:X81371; NID:9563340; PIDN:CAA57136.1; PID:9563341  
C:Superfamily: secretory component; immunoglobulin homology  
C:Keywords: duplication; glycoprotein; immunoglobulin receptor; phosphoprotein; tr-  
F:1-18/Domain: signal sequence #status predicted <Sig>  
F:19-757/Product: transmembrane secretory component #status predicted <MATM>  
F:33-112/Domain: immunoglobulin homology <IM1>  
F:145-222/Domain: immunoglobulin homology <IM2>  
F:250-328/Domain: immunoglobulin homology <IM3>  
F:365-444/Domain: immunoglobulin homology <IM4>  
F:476-547/Domain: immunoglobulin homology <IM5>  
F:631-653/Domain: transmembrane #status predicted <TMN>  
F:654-757/Domain: intracellular #status predicted <IN>  
F:40-110, 56-64, 152-220, 166-173, 257-374, 271-279, 370-440, 384-394, 481-543, 495-502/Dis-  
F:83, 420, 468/Binding site: carbohydrate (asn) (covalent) #status predicted  
F:665/Binding site: phosphate (Ser) (covalent) #status predicted

Query Match 2.7%: Score 21: DB 1: Length 757:  
Best Local Similarity 100.0%: Pred. No. 2,3e-12:  
Matches 21: Conservative 0: Mismatches 0: Indels 0: Gaps 0:

QY 40 CYPPTSVNRHTRKYWCROGA 60  
Db 40 CYPPTSVNRHTRKYWCROGA 60



## RESULT 3

145956  
polymeric immunoglobulin receptor - bovine  
C:Species: Bos primigenius taurus (cattle)  
C:Date: 19-Dec-1997 #sequence\_revision 19-Dec-1997 #text\_change 23-Jul-1999  
C:Accession: 145956  
R:Kuslert, M.A.; Krafc, P.; Myklebust, O.; Rognes, S.  
DNA Cell Biol. 14, 251-256, 1995  
A:Title: Cloning and characterization of two forms of bovine polymeric immunoglobulin re  
A:Reference number: 145956; MUID:95186063; PMID:7880445  
A:Accession: 145956  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-757 <KUL>  
A:Cross-references: GB:L04797; NID:9388279; PIDN:AC41620.1; PID:9388280  
C:Superfamily: secretory component; immunoglobulin homology  
F:145-222/Domain: immunoglobulin homology <IM1>

Query Match 2.7%; Score 21; DB 1; Length 757;  
Best Local Similarity 100.0%; Pred. No. 2,38-12;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 40 CYPTSVNRHRRKRCROGA 60  
|||||  
DB 40 CYPTSVNRHRRKRCROGA 60

## RESULT 4

## ORRGS

secretory component precursor - rat  
N:Alternate names: poly-Ig receptor; polymeric immunoglobulin receptor  
N:Contains: free secretory component; transmembrane secretory component  
C:Species: Rattus norvegicus (Norway rat)  
C:Date: 07-Sep-1990 #sequence\_revision 23-Aug-1996 #text\_change 16-Jul-1999  
C:Accession: S05407; S54731  
R:Bantling, G.; Brake, B.; Braghetta, P.; Luzzo, J.P.; Stanley, K.K.  
FEBS Lett. 254, 177-183, 1989  
A:Title: Intracellular targeting signals of polymeric immunoglobulin receptors are high  
A:Reference number: S05407; MUID:89378226; PMID:2776882  
A:Accession: S05407  
A:Molecule type: mRNA  
A:Residues: 1-769 <BAN>  
A:Cross-references: EMBL:X15741; NID:956464; PIDN:CAA3758.1; PID:956465  
C:Complex: monomeric as a transmembrane receptor or free in mucosal secretions; heterod  
amers; hetero-22-mer composed of one chain of secretory component, one chain of immunog  
C:Superfamily: secretory component; immunoglobulin homology  
C:Keywords: duplication; glycoprotein; immunoglobulin receptor; phosphoprotein; transcyt  
F:1-18/Domain: signal sequence #status predicted <SIG>  
F:19-769/Product: transmembrane secretory component #status predicted <MATM>  
F:19-582/Product: free secretory component #status predicted <MATF>  
F:33-112/Domain: immunoglobulin homology <IM1>  
F:145-222/Domain: immunoglobulin homology <IM2>  
F:250-326/Domain: immunoglobulin homology <IM3>  
F:363-442/Domain: immunoglobulin homology <IM4>  
F:477-548/Domain: immunoglobulin homology <IM5>  
F:644-666/Domain: transmembrane #status predicted <TM>  
F:667-769/Domain: intracellular #status predicted <INT>  
F:90-110,56-64,152-220,370-440,384-484,488-546,498-505/Disulfide bonds: #status predict  
F:90,135,471/Binding site: carbohydrate (asn) (covalent) #status predicted  
F:488-522/Disulfide bonds: (in Ig-unbound form) #status predicted  
F:488/Disulfide bonds: Interchain (to Iga alpha-1 chain-192) #status predicted  
F:522/Binding site: cysteine (cys) (covalent) (in Ig-bound form) #status predicted  
F:678/Binding site: phosphate (ser) (covalent) #status predicted

Query Match 2.2%; Score 17; DB 1; Length 769;  
Best Local Similarity 100.0%; Pred. No. 2,66-08;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 642 STLVPLGLVAVAV 658  
|||||  
DB 647 STLVPLGLVAVAV 663

## RESULT 5

## ORRBC

secretory component precursor - rabbit  
N:Alternate names: poly-Ig receptor; polymeric immunoglobulin receptor  
N:Contains: free secretory component; transmembrane secretory component  
C:Species: Oryctolagus cuniculus (domestic rabbit)  
C:Date: 15-Nov-1984 #sequence\_revision 15-Nov-1984 #text\_change 16-Jul-1999  
C:Accession: A02111; A28077  
R:Mostov, K.E.; Friedlander, M.; Blobel, G.  
Nature 308, 37-43, 1984

A:Title: The receptor for trans epithelial transport of Iga and Igm contains multi  
A:Reference number: A02111; MUID:84142246; PMID:6322002  
A:Accession: A02111  
A:Molecule type: mRNA  
A:Residues: 1-773 <MOS>  
A:Cross-references: GB:X00412; GB:K01291; NID:91595; PIDN:CAA25118.1; PID:91596  
A:Note: the authors translated the codon ACC for residue 54 as Asn  
R:Frutiger, S.; Hughes, G.J.; Hanly, W.C.; Jaton, J.C.  
J. Biol. Chem. 263, 8120-8125, 1988  
A:Title: Rabbit secretory components of different allotypes vary in their carbohy  
A:Reference number: A28077; MUID:88228032; PMID:3131339  
A:Accession: A28077

A:Molecule type: protein  
A:Residues: 87-114,410-424 <FRU>  
C:Comment: This receptor binds polymeric Iga and Igm at the basolateral surface o  
process, cleavage occurs to separate the extracellular portion, also known as the  
C:Comment: The five domains exhibit homology with immunoglobulin V regions. The s  
C:Comment: Alternative splicing in the extracellular domain leads to high or low  
C:Superfamily: secretory component; immunoglobulin homology  
C:Keywords: alternative splicing; duplication; glycoprotein; immunoglobulin recep  
F:1-18/Domain: signal sequence #status predicted <SIG>  
F:19-773/Product: transmembrane secretory component #status predicted <MATM>  
F:19-575/Product: free secretory component #status predicted <MATF>  
F:30-647/Domain: extracellular #status predicted <EXT>  
F:39-117/Domain: immunoglobulin homology <IM1>  
F:148-227/Domain: immunoglobulin homology <IM2>  
F:253-326/Domain: immunoglobulin homology <IM3>  
F:362-440/Domain: immunoglobulin homology <IM4>  
F:471-540/Domain: immunoglobulin homology <IM5>  
F:648-670/Domain: transmembrane #status predicted <TM>  
F:671-773/Domain: intracellular #status predicted <INT>  
F:46-115,155-225,260-324,369-438,478-538/Disulfide bonds: #status predicted  
F:108/Binding site: carbohydrate (asn) (covalent) (partial) #status experimental  
F:418/Binding site: carbohydrate (asn) (covalent) #status experimental

Query Match 1.7%; Score 13; DB 1; Length 773;  
Best Local Similarity 100.0%; Pred. No. 0,00029;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 727 EPRKAKRSKEA 739  
|||||  
DB 736 EPRKAKRSKEA 748

## RESULT 6

## B84311

hypothetical protein Vng1578h [imported] - Halobacterium sp. NRC-1  
C:Species: Halobacterium sp. NRC-1  
C:Date: 02-Feb-2001 #sequence\_revision 02-Feb-2001 #text\_change 16-Feb-2001  
C:Accession: B84311  
R:Ng, W.V.; Kennedy, S.P.; Mahairas, G.G.; Bergquist, B.; Pan, M.; Shukla, H.D.; L  
; Leitthausen, B.; Keller, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Maddocks, D.G.  
Jung, K.H.; Alam, M.; Prells, T.  
Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000  
A:Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ehardt, H.; Lowe, T  
A:Title: Genome sequence of Halobacterium species NRC-1  
A:Reference number: A84160; MUID:20504483; PMID:11016930  
A:Accession: B84311  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-308 <STO>  
A:Cross-references: GB:AE004437; NID:q10581062; PIDN:AG19854.1; GSPDB:GN00138  
C:Genetics:

A:Gene: VNG1578H  
C:Superfamily: cobalamin biosynthesis protein D

Query Match 1.38; Score 10; DB 2; Length 308;  
Best Local Similarity 100.0%; Pred. No. 0.15;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 652 AVGAVAVGVA 661  
|||||  
Db 5 AVGAVAVGVA 14

# RESULT 7

T35253

Small hypothetical protein SC5F2A.11 - Streptomyces coelicolor

C:Species: Streptomyces coelicolor  
C:Date: 05-Nov-1999 #sequence\_revision 05-Nov-1999 #text\_change 05-Nov-1999

C:Accession: T35253

R:Oliver, K.; Harris, D.; Bentley, S.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A. submitted to the EMBL Data Library, April 1999

A:Reference number: 221573

A:Accession: T35253

A:Status: preliminary; translated from GB/EMBL/DDBJ

A:Molecule type: DNA

A:Residues: 1-48 <OLIT>

A:Cross-references: EMBL:AL049587; PIDN:CAB40678.1; GSPDB:GN00070; SCOEDB:SC5F2A.11

A:Experimental source: strain A3(2)

C:Genetics:  
A:Gene: SCOEDB:SC5F2A.11

Query Match 1.0%; Score 8; DB 2; Length 48;  
Best Local Similarity 100.0%; Pred. No. 3.3;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 648 GLVLAVGA 655  
|||||  
Db 24 GLVLAVGA 31

# RESULT 8

E64402

hypothetical protein M0821 - Methanococcus jannaschii

C:Species: Methanococcus jannaschii  
C:Date: 13-Sep-1996 #sequence\_revision 13-Sep-1996 #text\_change 22-Oct-1999

C:Accession: E64402

R:Bult, C.J.; White, O.; Olsen, G.J.; Zhou, L.; Fleischmann, R.D.; Sutton, G.G.; Blake, R.; Reich, C.I.; Overbeek, R.; Kirkness, E.F.; Weissbrock, K.G.; Merrick, J.M.; Glodek, A.; Ison, J.D.; Sadow, P.W.; Hanna, M.C.; Cotton, M.D.; Roberts, K.M.; Hurst, M.A. Science 273, 1058-1073, 1996

A:Authors: Kaine, B.P.; Borodovsky, M.; Klenk, H.P.; Fraser, C.M.; Smith, H.O.; Woese, C.A.; Title: Complete genome sequence of the methanogenic archaeon, Methanococcus jannaschii

A:Reference number: A64300; M0ID:96337999; PMID:868087

A:Accession: E64402

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-56 <BLU>

A:Cross-references: GB:U67526; GB:L77117; NID:q1591508; PIDN:AB98833.1; PID:q1499645; T

C:Genetics:  
A:Map position: REV742767-742597

Query Match 1.0%; Score 8; DB 2; Length 56;  
Best Local Similarity 100.0%; Pred. No. 3.7;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 457 IIEGPNL 464  
|||||  
Db 16 IIEGPNL 23

# RESULT 9

F71248

probable proteasome beta subunit - Pyrococcus horikoshii

C:Species: Pyrococcus horikoshii

C:Date: 14-Aug-1998 #sequence\_revision 14-Aug-1998 #text\_change 21-Jul-2000

C:Accession: F71248

R:Kawarayashi, Y.; Sawada, M.; Horikawa, H.; Halkawa, Y.; Hino, Y.; Yamamoto, S.; M.; Ohfuku, Y.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; Kushiida, N.; DNA Res. 5, 55-76, 1998

A:Title: Complete sequence and gene organization of the genome of a hyper-thermoph

A:Reference number: A71000; M0ID:98344137; PMID:9679194

A:Accession: F71248

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-197 <KAM>

A:Cross-references: GB:AP000001; NID:q3236128; PIDN:BAA29317.1; PID:q3256634

A:Experimental source: strain OT3

A:Note: this accession replaces an interim accession for a sequence replaced by Ge

C:Genetics:  
A:Gene: PH0245

C:Superfamily: multicatalytic endopeptidase complex chain C9

Query Match 1.0%; Score 8; DB 2; Length 197;  
Best Local Similarity 100.0%; Pred. No. 11;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 566 KAASRDV 573  
|||||  
Db 166 KAASRDV 173

# RESULT 10

E90206

hypothetical protein hisf [imported] - Sulfolobus solfataricus

C:Species: Sulfolobus solfataricus

C:Date: 24-May-2001 #sequence\_revision 24-May-2001 #text\_change 15-Jun-2001

C:Accession: E90206

R:She, Q.; Singh, R.K.; Confalonieri, F.; Zivanovic, Y.; Allard, G.; Awayez, M.J.; aretti, R.A.; Ragan, M.A.; Sengen, C.W.; Van der Oost, J. submitted to Genbank, April 2001

A:Description: Sulfolobus solfataricus complete genome.

A:Reference number: A99139

A:Accession: E90206

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-251 <KUR>

A:Cross-references: GB:AE006641; NID:q13813762; PIDN:AAK40908.1; GSPDB:GN00155

C:Genetics:  
A:Gene: hisf

C:Superfamily: cyclase hisf

Query Match 1.0%; Score 8; DB 2; Length 251;  
Best Local Similarity 100.0%; Pred. No. 13;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 574 SLAKADA 581  
|||||  
Db 215 SLAKADA 222

RESULT 11  
S55662  
legument protein 67 - equine herpesvirus 2.  
C:Species: equine herpesvirus 2  
C:Date: 27-Oct-1995 #sequence\_revision 03-Nov-1995 #text\_change 26-Aug-1999

C:Accession: S55662  
R:Telford, E.A.R.; Watson, M.S.; Auld, H.C.; Perry, J.; Davison, A.J. J. Mol. Biol. 249, 520-528, 1995  
A:Title: The DNA sequence of equine herpesvirus 2.  
A:Reference number: S55594; M0ID:95302501; PMID:7783207  
A:Accession: S55662  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-287 <REL>  
A:Cross-references: GB:U20824; NID:q695172; PIDN:AAK13855.1; PID:q695240  
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, February 1-

C:Superfamily: human herpesvirus 4 BFRF1 protein

Query Match 1.0%; Score 8; DB 2; Length 287;  
Best Local Similarity 100.0%; Pred. No. 15;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 650 VLAVGAVA 657  
Db 268 VLAVGAVA 275

RESULT 12

conserved hypothetical protein VC0480 [Imported] - Vibrio cholerae (strain N16961 serog  
C:Species: Vibrio cholerae  
C:Date: 18-Aug-2000 #sequence\_revision 20-Aug-2000 #text\_change 02-Feb-2001

C:Accession: A82318  
R:Heidelberg, J.F.; Elsen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwin, M.L.; Dodson, R.J.;  
Charlston, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragol, I.; Sellers, P.  
1, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.  
Nature 406, 477-483, 2000

A:Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.  
A:Reference number: A82035; MUID:20406833; PMID:10952301

A:Accession: A82318  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-287 <HEI>  
A:Cross-references: GB:AE004134; GB:AE003852; NID:9654900; PIDN:APF93653.1; GSPDB:GN001  
C:Experimental source: serogroup O1; strain N16961; biotype El Tor  
C:Genetics:  
A:Gene: VC0480  
A:Map position: 1  
C:Superfamily: Escherichia coli hypothetical 30.9K protein (sbm-fba intergenic region)

Query Match 1.0%; Score 8; DB 2; Length 287;  
Best Local Similarity 100.0%; Pred. No. 15;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 358 VVKGVACS 365  
Db 50 VVKGVACS 57

RESULT 13

T45133  
restriction endonuclease Mami [Imported] - Microbacterium ammoniaphilum

N:Alternate names: restriction-modification system Mami

C:Species: Microbacterium ammoniaphilum

C:Date: 21-Jan-2000 #sequence\_revision 21-Jan-2000 #text\_change 16-Feb-2001

C:Accession: T45133

R:Striebel, H.M.; Seeber, S.; Jarsch, M.; Kessler, C.  
Gene 172, 41-46, 1996

A:Title: Cloning and characterization of the Mami restriction-modification system from M

A:Reference number: Z22923; MUID:96257250; PMID:8654988

A:Accession: T45133

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-308 <STR>

A:Cross-references: EMBL:X79027; NID:9984667; PIDN:CA55648.1; PID:9984669

A:Experimental source: ATCC 15354

C:Genetics:

A:Gene: MamiR

C:Superfamily: Microbacterium ammoniaphilum restriction endonuclease Mami

Query Match 1.0%; Score 8; DB 2; Length 308;  
Best Local Similarity 100.0%; Pred. No. 16;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 280 DVVNTLG 287  
Db 200 DVVNTLG 207

RESULT 14

D64462

hypothetical protein MJ1301 - Methanococcus jannaschii

C:Species: Methanococcus jannaschii

C:Date: 13-Sep-1996 #sequence\_revision 13-Sep-1996 #text\_change 21-Jul-2000

C:Accession: D64462

R:Bult, C.J.; White, O.; Olsen, G.J.; Zhou, L.; Fleischmann, R.D.; Sutton, G.G.; I

Reich, C.I.; Overbeek, R.; Kirsch, E.F.; Weinstock, K.G.; Merrick, J.M.; Glode

rson, J.D.; Sadow, P.W.; Hanna, M.C.; Cotton, M.D.; Roberts, K.M.; Hurst, M.A.

Science 273, 1058-1073, 1996

A:Authors: Kaine, B.P.; Borodovsky, M.; Klenk, H.P.; Fraser, C.M.; Smith, H.O.; Wk

A:Title: Complete genome sequence of the methanogenic archaeon, Methanococcus jani

A:Reference number: A64300; MUID:96337999; PMID:8688087

A:Accession: D64462

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-400 <BUL>

A:Cross-references: GB:U67571; GB:L77117; NID:q1591939; PIDN:AAB99326.1; PID:q159;

C:Genetics:

A:Map position: REV1250122-1248920

C:Superfamily: ATP-binding protein PAB1945

Query Match 1.0%; Score 8; DB 2; Length 400;  
Best Local Similarity 100.0%; Pred. No. 19;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 457 IIEGEPNL 464  
Db 23 IIEGEPNL 30

RESULT 15

T21834

hypothetical protein F36A2.4 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 15-Oct-1999

C:Accession: T21834

R:Lennard, N.  
submitted to the EMBL Data Library, October 1996

A:Reference number: Z19476

A:Accession: T21834

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-528 <WIL>

A:Cross-references: EMBL:Z81077; PIDN:CAB03071.1; GSPDB:GN00019; CESP:F36A2.4

A:Experimental source: Clone F36A2

C:Genetics:

A:Gene: CESP:F36A2.4

A:Map position: 1

A:Introns: 45/3; 81/3; 114/3; 151/3; 195/1; 306/1; 326/1; 368/2; 413/1; 464/2; 494

Query Match 1.0%; Score 8; DB 2; Length 528;  
Best Local Similarity 100.0%; Pred. No. 25;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 263 VANYAKFL 270  
Db 203 VANYAKFL 210

Search completed: September 9, 2003, 23:18:39  
Job time : 47 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 9, 2003, 23:07:32 : Search time 26 Seconds

(Without alignments)  
1381.862 Million cell updates/sec

Title: US-09-818-247-1

Perfect score: 764  
Sequence: 1 MLFVLTCLLAVPAISTKS.....DFLLQSTVAEAADGPOEA 764

Scoring table: OLIGO

Searched: 127863 seqs, 47026705 residues

Word size : 0

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database : SwissProt\_41.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	764	100.0	764	1	PIGR_HUMAN
2	21	2.7	757	1	PIGR_BOVIN
3	17	2.2	769	1	PIGR_RAT
4	16	2.1	771	1	PIGR_MOUSE
5	13	1.7	773	1	PIGR_RABIT
6	11	1.4	102	1	PIGR_PIG
7	8	1.0	56	1	Y821_METJA
8	8	1.0	251	1	HIS6_SULSO
9	8	1.0	308	1	T2M1_MICAM
10	8	1.0	400	1	YD01_METJA
11	8	1.0	556	1	SYFB_PYRAB
12	8	1.0	556	1	SYFB_PYRAB
13	8	1.0	563	1	B2NB_PSEFL
14	8	1.0	802	1	FCR4_HUMAN
15	8	1.0	808	1	FCR4_MOUSE
16	8	1.0	1298	1	VGR3_HUMAN
17	8	1.0	1363	1	VGR3_MOUSE
18	8	1.0	1712	1	NX2A_HUMAN
19	8	1.0	1715	1	NX2A_RAT
20	7	0.9	94	1	YF8_YEAST
21	7	0.9	118	1	Y8P1_STRCU
22	7	0.9	120	1	Y025_MYCTU
23	7	0.9	127	1	R22A_ARATH
24	7	0.9	128	1	WNT8_THUTH
25	7	0.9	128	1	YF70_AQUAE
26	7	0.9	143	1	CAS2_MOUSE
27	7	0.9	153	1	HXC6_SHEEP
28	7	0.9	163	1	IPYR_STRCO
29	7	0.9	209	1	PGPI_HUMAN
30	7	0.9	209	1	PGPI_MOUSE
31	7	0.9	228	1	MTH2_HAENA
32	7	0.9	233	1	US56_HSVIE
33	7	0.9	233	1	HXC6_MOUSE

34	7	0.9	247	1	HXA4_HETFR	091a22 heterodontu
35	7	0.9	250	1	COX2_NEUCR	P00411 neurospora
36	7	0.9	255	1	HXD4_HUMAN	P09016 homo sapien
37	7	0.9	257	1	MOTA_HELPY	P56426 heliobacte
38	7	0.9	259	1	YGFW_ECOLI	046813 escherichia
39	7	0.9	277	1	KD5A_BROBE	08fh1 brucella me
40	7	0.9	277	1	KD5A_RHILU	08m26 rhizobium l
41	7	0.9	278	1	Y110_MYCPN	P75523 mycoplasma
42	7	0.9	279	1	H7PX_NEIMB	09K006 neisseria m
43	7	0.9	280	1	KD5A_RHIME	092q99 rhizobium m
44	7	0.9	282	1	KD5A_AGRF5	08ufh3 agrobacteri
45	7	0.9	285	1	SYGB_HAENA	030836 haemophilus

## ALIGNMENTS

RESULT 1  
PIGR\_HUMAN  
ID PIGR\_HUMAN STANDARD: PRT: 764 AA.  
AC P01833;  
DT 21-JUL-1986 (Rel. 01, Created)  
DT 01-OCT-1996 (Rel. 34, Last sequence update)  
DT 15-SEP-2003 (Rel. 42, Last annotation update)  
DE Polymeric-immunoglobulin receptor precursor (Poly-Ig receptor) (PIGR)  
DE [Contains: Secretory component].  
GN PIGR.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=92039621; PubMed=1682231;  
RA Brandtzaeg P.; Grzeschik K.H.; Geurts van Kessel A.H.; Olaisen B.,  
Krajacl P.; Solberg R.; Sandberg M.; Oyen O.; Jahnson T.,  
Brandtzaeg P.; "The human transmembrane secretory component (poly-Ig receptor):  
RT molecular cloning, restriction fragment length polymorphism and  
RT chromosomal sublocalization.";  
RL Hum. Genet. 87:642-648(1991).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=92387236; PubMed=1355431;  
RA Krajacl P.; Kvale D.; Tasken K.; Brandtzaeg P.;  
RT "Molecular cloning and exon-intron mapping of the gene encoding human  
RT transmembrane secretory component (the poly-Ig receptor).";  
RL Eur. J. Immunol. 22:2309-2315(1992).  
RN [3]  
RP SEQUENCE OF 72-764 FROM N.A.  
RX MEDLINE=89149795; PubMed=2920039;  
RA Krajacl P.; Solberg R.; Sandberg M.; Oyen O.; Jahnson T.,  
Brandtzaeg P.; "Molecular cloning of the human transmembrane secretory component  
RT (poly-Ig receptor) and its mRNA expression in human tissues.";  
RL Biochem. Biophys. Res. Commun. 156:783-789(1989).  
RN [4]  
RP SEQUENCE OF 19-577, DISULFIDE BONDS, AND CARBOHYDRATE-LINKAGE SITES.  
RX MEDLINE=85128981; PubMed=6526384;  
RA Eifert H.; Quentlin E.; Decker J.; Hillemeir S.; Hufschmidt M.,  
Klingmüller D.; Weber M.H.; Hilschmann N.;  
RT "The primary structure of human free secretory component and the  
RT arrangement of disulfide bonds.";  
RL Hoppe-Seyler's Z. Physiol. Chem. 365:1489-1495(1984).  
RN [5]  
RP SEQUENCE OF 19-577.  
RX MEDLINE=91315750; PubMed=1859628;  
RA Eifert H.; Quentlin E.; Wiederhold M.; Hillemeir S.; Decker J.,  
Weber M.; Hilschmann N.;  
RT "Determination of the molecular structure of the human free secretory  
RT component.";  
RL Biol. Chem. Hoppe-Seyler 372:119-126(1991).  
RN [6]  
RP SEQUENCE OF 118-138; 212-230; 232-268; 273-288 AND 578-603.

RX MEDLINE-97379357; PubMed-9237679;  
 RA Hughes G.J., Fritzliger S., Savoy L.-A., Reason A.J., Morris H.R.,  
 RA Jaton J.-C.;  
 RT "Human free secretory component is composed of the first 585 amino  
 RT acid residues of the polymeric immunoglobulin receptor.";  
 RL FEBS Lett. 410:443-446(1997).  
 CC -!- FUNCTION: THIS RECEPTOR BINDS POLYMERIC IGA AND IGM AT THE  
 CC BASOLATERAL SURFACE OF EPITHELIAL CELLS. THE COMPLEX IS THEN  
 CC TRANSPORTED ACROSS THE CELL TO BE SECRETED AT THE APICAL SURFACE.  
 CC DURING THIS PROCESS A CLEAVAGE OCCURS THAT SEPARATE THE  
 CC EXTRACELLULAR (KNOWN AS THE SECRETORY COMPONENT) FROM THE  
 CC TRANSMEMBRANE SEGMENT.  
 CC -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN. ALSO SECRETED.  
 CC -!- SIMILARITY: Contains 5 immunoglobulin-like V-type domains.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: S62403; AAB20203.1; -;  
 DR EMBL: S43443; AAB23176.1; -;  
 DR EMBL: S43437; AAB23176.1; JOINED.  
 DR EMBL: S43441; AAB23176.1; JOINED.  
 DR EMBL: S43442; AAB23176.1; JOINED.  
 DR EMBL: S43443; AAB23176.1; JOINED.  
 DR EMBL: S43444; AAB23176.1; JOINED.  
 DR EMBL: S43445; AAB23176.1; JOINED.  
 DR EMBL: S43446; AAB23176.1; JOINED.  
 DR EMBL: S43447; AAB23176.1; JOINED.  
 DR EMBL: S43448; AAB23176.1; JOINED.  
 DR EMBL: M24559; AAA36102.1; -;  
 DR EMBL: A52091; CAA03384.1; -;  
 DR PIR: A46537; QRHUGS;  
 DR GlycosultedB; P01833; -;  
 DR Genew; HGNC:8968; PIGR.  
 DR MIM: 173880; -;  
 DR GO: GO:0005887; C: integral to plasma membrane; TAS.  
 DR InterPro: IPR007110; Ig-like.  
 DR InterPro: IPR003599; Ig.  
 DR InterPro: IPR003006; Ig\_MHC.  
 DR Pfam: PF00047; Ig; 5.  
 DR SMART: SM00409; IG; 5.  
 DR PROSITE: PS50835; IG LIKE; 2.  
 KW Immunoglobulin domain; Repeat; Transmembrane; Glycoprotein; Signal;  
 KW Polymorphism.  
 FT SIGNAL 1 18  
 FT CHAIN 19 764 POLYMERIC-IMMUNOGLOBULIN RECEPTOR.  
 FT CHAIN 19 603 SECRETORY COMPONENT.  
 FT DOMAIN 19 638 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 639 661 POTENTIAL.  
 FT DOMAIN 662 764 CYTOPLASMIC (POTENTIAL).  
 FT DOMAIN 19 120 IG-LIKE V-TYPE 1.  
 FT DOMAIN 145 237 IG-LIKE V-TYPE 2.  
 FT DOMAIN 250 352 IG-LIKE V-TYPE 3.  
 FT DOMAIN 364 458 IG-LIKE V-TYPE 4.  
 FT DOMAIN 462 561 IG-LIKE V-TYPE 5.  
 FT DISULFID 40 110  
 FT DISULFID 56 64  
 FT DISULFID 152 220  
 FT DISULFID 257 325  
 FT DISULFID 271 279  
 FT DISULFID 371 441  
 FT DISULFID 385 393  
 FT DISULFID 482 544  
 FT DISULFID 486 520  
 FT DISULFID 496 503  
 FT CARBOHYD 83 83  
 FT CARBOHYD 90 90  
 FT CARBOHYD 135 135  
 N-LINKED (GLCNAC. . .)  
 N-LINKED (GLCNAC. . .)  
 N-LINKED (GLCNAC. . .)  
 N-LINKED (GLCNAC. . .)

FT CARBOHYD 186 186 N-LINKED (GLCNAC. . .)  
 FT CARBOHYD 421 421 N-LINKED (GLCNAC. . .)  
 FT CARBOHYD 469 469 N-LINKED (GLCNAC. . .)  
 FT CARBOHYD 499 499 N-LINKED (GLCNAC. . .)  
 FT VARIANT 580 580 A -> V.  
 FT VARIANT 580 580 /PRTD-VAR. 003920.  
 FT CONFLICT 136 136 D -> Q (IN REF. 4 AND 5).  
 FT CONFLICT 158 158 N -> D (IN REF. 4 AND 5).  
 FT CONFLICT 208 209 NO -> DE (IN REF. 4 AND 5).  
 FT CONFLICT 229 229 MISSING (IN REF. 4 AND 5).  
 FT CONFLICT 234 234 D -> N (IN REF. 4 AND 5).  
 FT CONFLICT 241 241 E -> Q (IN REF. 4 AND 5).  
 FT CONFLICT 262 262 E -> Q (IN REF. 4 AND 5).  
 FT CONFLICT 280 280 D -> N (IN REF. 4 AND 5).  
 FT CONFLICT 392 392 N -> D (IN REF. 4 AND 5).  
 FT CONFLICT 500 500 N -> D (IN REF. 4 AND 5).  
 SQ SEQUENCE 764 AA; 83313 MW; 916B3E662C339950 CRC64;  
 Query Match 100.0%; Score 764; DB 1; Length 764;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 764; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 MLFVLTCLAVPAISTKSPFGPEVNSVSGNSVSYNCYPTSVNRRTRKYMGROGA 60  
 DB 1 MLFVLTCLAVPAISTKSPFGPEVNSVSGNSVSYNCYPTSVNRRTRKYMGROGA 60  
 QY 61 RGGCITLISSEGYVSSKYAGRANLJTFPENGTFVYVIAQLSQDSCGRYKCGIGNSRGLS 120  
 DB 61 RGGCITLISSEGYVSSKYAGRANLJTFPENGTFVYVIAQLSQDSCGRYKCGIGNSRGLS 120  
 QY 121 FDVLSVSGPGGLNLTQKTYTVLQGRVYINCFFKRNMAKRSYKQJGLYVLYVLDSS 180  
 DB 121 FDVLSVSGPGGLNLTQKTYTVLQGRVYINCFFKRNMAKRSYKQJGLYVLYVLDSS 180  
 QY 181 GYVNPYVYGRIRLDIGTGGLFSVYINQLRLSDAGQYLCQAGDSDSNKKNADLYLKP 240  
 DB 181 GYVNPYVYGRIRLDIGTGGLFSVYINQLRLSDAGQYLCQAGDSDSNKKNADLYLKP 240  
 QY 241 EPELVYEDLGSVTFHCALGPEVANYAKFLCROSSSENCDDVYVNLGKRAPAFEGRIILN 300  
 DB 241 EPELVYEDLGSVTFHCALGPEVANYAKFLCROSSSENCDDVYVNLGKRAPAFEGRIILN 300  
 QY 301 POKKDSFSVYITGLKREDAGRYLCGAHSGQLOESPIQAMOLFVNESTIPRSPYVK 360  
 DB 301 POKKDSFSVYITGLKREDAGRYLCGAHSGQLOESPIQAMOLFVNESTIPRSPYVK 360  
 QY 361 GVAQSSVAVLCPRNKRRESKISKYKCLMEGANGRCPLVDSBGMVKAQYGRSLSEEPG 420  
 DB 361 GVAQSSVAVLCPRNKRRESKISKYKCLMEGANGRCPLVDSBGMVKAQYGRSLSEEPG 420  
 QY 421 NGFTVYLNQLTSRDAGFYWCLTNGDTLMFTVEIKIIEGPNLKYPGNVTAVLGFTLVY 480  
 DB 421 NGFTVYLNQLTSRDAGFYWCLTNGDTLMFTVEIKIIEGPNLKYPGNVTAVLGFTLVY 480  
 QY 481 PCGFPECKFSSEYKWKCKNNNTGQALPQSDGSPSKAFVNCNDENSRVLSLTNLVTRADEG 540  
 DB 481 PCGFPECKFSSEYKWKCKNNNTGQALPQSDGSPSKAFVNCNDENSRVLSLTNLVTRADEG 540  
 QY 541 WYMGYKQGFYGETAAVYVAVEERKAAGSRDYSLAKADAPPEKYVLDGSGFREIENKAIO 600  
 DB 541 WYMGYKQGFYGETAAVYVAVEERKAAGSRDYSLAKADAPPEKYVLDGSGFREIENKAIO 600  
 QY 601 DPLFAEERKAVADTRDQADGSRASVDSGSEEDGSSRALVSTLVPLGLVLAAGAAGVY 660  
 DB 601 DPLFAEERKAVADTRDQADGSRASVDSGSEEDGSSRALVSTLVPLGLVLAAGAAGVY 660  
 QY 661 ARAHRKNDVRSVIRSTRDISDSDENSRFGANDNMGASSTTOETSLGCKEFAATTE 720  
 DB 661 ARAHRKNDVRSVIRSTRDISDSDENSRFGANDNMGASSTTOETSLGCKEFAATTE 720  
 QY 721 STTETPEPKAKRSSKEEAEMAYKDFLQSSYVAAEADQDPOBA 764  
 DB 721 STTETPEPKAKRSSKEEAEMAYKDFLQSSYVAAEADQDPOBA 764

FT	CHAIN	19	757	POLYMERIC-IMMUNOGLOBULIN RECEPTOR.
FT	CHAIN	19	599	SECRETORY COMPONENT (BY SIMILARITY).
FT	DOMAIN	19	632	EXTRACELLULAR (POTENTIAL).
FT	TRANSMEM	633	653	POTENTIAL.
FT	DOMAIN	654	757	CYTOPLASMIC (POTENTIAL).
FT	DOMAIN	19	126	IG-LIKE V-TYPE 1.
FT	DOMAIN	145	237	IG-LIKE V-TYPE 2.
FT	DOMAIN	250	341	IG-LIKE V-TYPE 3.
FT	DOMAIN	353	457	IG-LIKE V-TYPE 4.
FT	DOMAIN	461	560	IG-LIKE V-TYPE 5.
FT	DISULFID	40	110	BY SIMILARITY.
FT	DISULFID	56	64	BY SIMILARITY.
FT	DISULFID	152	220	BY SIMILARITY.
FT	DISULFID	257	324	BY SIMILARITY.
FT	DISULFID	271	279	BY SIMILARITY.
FT	DISULFID	370	440	BY SIMILARITY.
FT	DISULFID	384	394	BY SIMILARITY.
FT	DISULFID	481	543	BY SIMILARITY.
FT	DISULFID	485	519	BY SIMILARITY.
FT	DISULFID	495	502	BY SIMILARITY.
FT	CARBOHYD	83	83	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	420	420	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	468	468	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	MOD RES	727	727	PHOSPHORYLATION (BY SIMILARITY).
FT	VARSPIC	129	346	Missing (in isoform short).
FT	VARSPIC	129	346	/FtId-VSP-002547.
FT	VARIANT	29	29	T -> S.
FT	VARIANT	142	142	V -> I.
FT	VARIANT	404	404	I -> M.
FT	VARIANT	413	413	A -> V.
FT	VARIANT	435	435	T -> A.
SO	SEQUENCE	757 AA;	82434 MW;	DCED67PDD6A6E6C6 CRC64;
Query Match 2.7%; Score 21; DB 1; Length 757;				
Best Local Similarity 100.0%; Pred. No. 5.2e-13;				
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;				
OY	40	CYPPTSVNRRTRKRYCROGA	60	
OY	40	CYPPTSVNRRTRKRYCROGA	60	
Db	40	CYPPTSVNRRTRKRYCROGA	60	
RESULT 3				
PIGR_RAT	PIGR_RAT	STANDARD;	PRT;	769 AA.
AC	P15083;			
DT	01-APR-1990 (Rel. 14, Created)			
DT	01-APR-1990 (Rel. 14, Last sequence update)			
DT	15-SEP-2003 (Rel. 42, Last annotation update)			
DE	Polymeric-immunoglobulin receptor precursor (Poly-Ig receptor) (PIGR)			
DE	[Contains: Secretory component].			
GN	PIGR.			
OS	Rattus norvegicus (Rat).			
OS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.			
OX	NCBI_TaxID=10116;			
RP	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Liver.			
RX	MEDLINE=89378226; PubMed=2776882;			
RA	Bailling G., Brake B., Braghetta P., Luzzo J.P., Stanley K.K.;			
RT	"Intracellular targeting signals of polymeric immunoglobulin			
RT	receptors are highly conserved between species.";			
RL	FEBS Lett. 254:177-183(1989).			
CC	-1- FUNCTION: THIS RECEPTOR BINDS POLYMERIC IGA AND IGM AT THE			
CC	BASOLATERAL SURFACE OF EPITHELIAL CELLS. THE COMPLEX IS THEN			
CC	TRANSPORTED ACROSS THE CELL TO BE SECRETED AT THE APICAL SURFACE.			
CC	DURING THIS PROCESS A CLEAVAGE OCCURS THAT SEPARATE THE			
CC	EXTRACELLULAR (KNOWN AS THE SECRETORY COMPONENT) FROM THE			
CC	TRANSMEMBRANE SEGMENT.			
CC	-1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN. ALSO SECRETED.			
CC	-1- SIMILARITY: Contains 5 immunoglobulin-like V-type domains.			
CC	-----			

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

EMBL: X15741; CAA33758.1; -  
 DR PIR: S05407; ORRGS.  
 DR InterPro: IPR007110; Ig-Like.  
 DR InterPro: IPR003599; Ig.  
 DR InterPro: IPR003006; Ig\_MHC.  
 DR Pfam: PF00047; Ig; 5.  
 DR SMART: SM00409; Ig; 5.  
 DR PROSITE: PS50835; IG-LIKE; 2.  
 DR Immunoglobulin domain; Repeat: Transmembrane; Glycoprotein; Signal;  
 FT SIGNAL 1 18  
 FT CHAIN 19 769 POLYMERIC-IMMUNOGLOBULIN RECEPTOR.  
 FT CHAIN 19 611 SECRETORY COMPONENT.  
 FT DOMAIN 19 643 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 644 666 POTENTIAL.  
 FT CYTOPLASM 667 769 CYTOPLASMIC (POTENTIAL).  
 FT DOMAIN 21 126 IG-LIKE V-TYPE 1.  
 FT DOMAIN 135 237 IG-LIKE V-TYPE 2.  
 FT DOMAIN 240 341 IG-LIKE V-TYPE 3.  
 FT DOMAIN 353 457 IG-LIKE V-TYPE 4.  
 FT DOMAIN 463 563 IG-LIKE V-TYPE 5.  
 FT DISULFID 40 110 POTENTIAL.  
 FT DISULFID 152 220 POTENTIAL.  
 FT DISULFID 257 324 POTENTIAL.  
 FT DISULFID 370 440 POTENTIAL.  
 FT CARBOHYD 90 90 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 135 135 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 206 206 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 471 471 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 SQ SEQUENCE 769 AA; 84798 MW; 5F84930340255A7 CRC64;

Query Match 2.2%; Score 17; DB 1; Length 769;  
 Best Local Similarity 100.0%; Pred. No. 7e-09;  
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 642 STVPLGLVLAAGAV 658  
 Db 647 STVPLGLVLAAGAV 663

RESULT 4  
 PIGR\_MOUSE STANDARD: PRT: 771 AA.  
 ID PIGR\_MOUSE  
 AC 070570;  
 DT 30-MAY-2000 (Rel. 39, Created)  
 DT 30-MAY-2000 (Rel. 39, Last sequence update)  
 DT 15-SEP-2003 (Rel. 42, Last annotation update)  
 DE PolymERIC-Immunoglobulin receptor precursor (Poly-Ig receptor) (PIGR)  
 DE [Contains: Secretory component].  
 GN PIGR.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN (1)  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6 X CBA; TISSUE=Liver;  
 RX MEDLINE=95138517; PubMed=7836758;  
 RA Pliskulich J.F., Blanchard M.H., Youngman K.R., France J.A.,  
 RA Kaetzel C.S.;  
 RA "Molecular cloning of the mouse polymeric Ig receptor. Functional  
 RA regions of the molecule are conserved among five mammalian species.";  
 RL J. Immunol. 154:1735-1747(1995).  
 RN (2)  
 RP SEQUENCE FROM N.A.

RC STRAIN=129/SVJ;  
 RX MEDLINE=98072444; PubMed=9409786;  
 RA Martin M.G., Gutierrez E.M., Lam J.T., Li T.W.H., Wang J.;  
 RA "Genomic cloning and structural analysis of the murine polymeric  
 RA receptor (PIGR) gene and promoter region.";  
 RL Gene 201:189-197(1997).  
 RN (3)  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=129; TISSUE=Liver;  
 RA de Groet N., Vollebregt E., Lee S.H., Verbeet M.P., de Boer H.A.;  
 RA "Molecular cloning and exon-intron organization of the gene encoding  
 RA the murine polymeric immunoglobulin receptor.";  
 RL Submitted (Feb-1996) to the EMBL/GenBank/DBJ databases.  
 CC -1- FUNCTION: THIS RECEPTOR BINDS POLYMERIC IGA AND IGM AT THE  
 CC BASOLATERAL SURFACE OF EPITHELIAL CELLS. THE COMPLEX IS THEN  
 CC TRANSPORTED ACROSS THE CELL TO BE SECRETED AT THE APICAL SURFACE.  
 CC DURING THIS PROCESS A CLEAVAGE OCCURS THAT SEPARATE THE  
 CC EXTRACELLULAR (KNOWN AS THE SECRETORY COMPONENT) FROM THE  
 CC TRANSMEMBRANE SEGMENT (BY SIMILARITY).  
 CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN. ALSO SECRETED.  
 CC -1- SIMILARITY: Contains 5 Immunoglobulin-like V-type domains.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

EMBL: U06431; AAA67440.1; -  
 DR EMBL: U83434; AAC53585.1; JOINED.  
 DR EMBL: U83427; AAC53585.1; JOINED.  
 DR EMBL: U83428; AAC53585.1; JOINED.  
 DR EMBL: U83429; AAC53585.1; JOINED.  
 DR EMBL: U83430; AAC53585.1; JOINED.  
 DR EMBL: U83431; AAC53585.1; JOINED.  
 DR EMBL: U83432; AAC53585.1; JOINED.  
 DR EMBL: U83433; AAC53585.1; JOINED.  
 DR EMBL: Y16524; CAA76272.1; JOINED.  
 DR EMBL: Y16525; CAA76272.1; JOINED.  
 DR EMBL: Y16526; CAA76272.1; JOINED.  
 DR EMBL: Y16527; CAA76272.1; JOINED.  
 DR EMBL: Y16528; CAA76272.1; JOINED.  
 DR EMBL: Y16529; CAA76272.1; JOINED.  
 DR EMBL: Y16530; CAA76272.1; JOINED.  
 DR EMBL: Y16531; CAA76272.1; JOINED.  
 DR EMBL: Y16532; CAA76272.1; JOINED.  
 DR MGD: MGI:103080; PIGR.  
 DR InterPro: IPR007110; Ig-Like.  
 DR InterPro: IPR003599; Ig.  
 DR InterPro: IPR003006; Ig\_MHC.  
 DR Pfam: PF00047; Ig; 5.  
 DR SMART: SM00409; Ig; 5.  
 DR PROSITE: PS50835; IG-LIKE; 3.  
 DR Immunoglobulin domain; Repeat: Transmembrane; Glycoprotein; Signal;  
 FT SIGNAL 1 18  
 FT CHAIN 19 771  
 FT CHAIN 19 611 POLYMERIC-IMMUNOGLOBULIN RECEPTOR.  
 FT DOMAIN 19 645 SECRETORY COMPONENT.  
 FT TRANSMEM 646 668 EXTRACELLULAR (POTENTIAL).  
 FT CYTOPLASM 669 771 POTENTIAL.  
 FT DOMAIN 21 120 IG-LIKE V-TYPE 1.  
 FT DOMAIN 135 237 IG-LIKE V-TYPE 2.  
 FT DOMAIN 245 351 IG-LIKE V-TYPE 3.  
 FT DOMAIN 352 457 IG-LIKE V-TYPE 4.  
 FT DOMAIN 463 563 IG-LIKE V-TYPE 5.  
 FT DISULFID 40 110 POTENTIAL.  
 FT DISULFID 152 220 POTENTIAL.  
 FT DISULFID 257 324 POTENTIAL.  
 FT DISULFID 370 440 POTENTIAL.  
 FT DISULFID 484 546 POTENTIAL.  
 FT CARBOHYD 90 90 N-LINKED (GLCNAC. . .) (POTENTIAL).

```

FT CARBOHYD 147 147 N-LINKED (GLCNAc. . .) (POTENTIAL).
FT CARBOHYD 170 170 N-LINKED (GLCNAc. . .) (POTENTIAL).
FT CARBOHYD 206 206 N-LINKED (GLCNAc. . .) (POTENTIAL).
FT CARBOHYD 420 420 N-LINKED (GLCNAc. . .) (POTENTIAL).
FT CARBOHYD 471 471 N-LINKED (GLCNAc. . .) (POTENTIAL).
FT CONFLICT 159 159 A -> V (IN REF. 1).
FT CONFLICT 396 396 V -> A (IN REF. 1).
FT CONFLICT 620 620 G -> R (IN REF. 1).
SQ SEQUENCE 771 AA; 84598 MW; 78C81302EC710730 CRC64;

Query Match 2.1%; Score 16; DB 1; Length 771;
Best Local Similarity 100.0%; Pred. No. 7.6e-08;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 45 TSVNRHRRKWCROGA 60
    |||||
Db 45 TSVNRHRRKWCROGA 60

RESULT 5
PIGR_RABIT STANDARD; PRT: 773 AA.
ID PIGR_RABIT STANDARD; PRT: 773 AA.
AC P01812;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Polymeric-immunoglobulin receptor precursor (Poly-Ig receptor) (PIGR)
DE [Contains: Secretory component].
GN PIGR.
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=84142246; PubMed=6322002;
RA Mostov K.E., Friedlander M., Blobel G.;
RT "The receptor for transepithelial transport of IgA and Igm contains
RT multiple immunoglobulin-like domains.";
RL Nature 308:37-43(1984).
RN [2]
RP SEQUENCE OF 87-114 AND 410-428.
RX MEDLINE=8828032; PubMed=313139;
RA Frutiger S., Hughes G.J., Hanly W.C., Jaton J.-C.;
RT "Rabbit secretory components of different allotypes vary in their
RT carbohydrate content and their sites of N-linked glycosylation.";
RL J. Biol. Chem. 263:8120-8125(1988).
CC -1- FUNCTION: THIS RECEPTOR BINDS POLYMERIC IGA AND IGM AT THE
CC BASOLATERAL SURFACE OF EPITHELIAL CELLS. THE COMPLEX IS THEN
CC TRANSPORTED ACROSS THE CELL TO BE SECRETED AT THE APICAL SURFACE.
CC DURING THIS PROCESS A CLEAVAGE OCCURS THAT SEPARATE THE
CC EXTRACELLULAR (KNOWN AS THE SECRETORY COMPONENT) FROM THE
CC TRANSMEMBRANE SEGMENT.
CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN. ALSO SECRETED.
CC -1- POLYMORPHISM: THE SEQUENCE SHOWN IS THAT OF ALLOTYPIC T62.
CC -1- SIMILARITY: Contains 5 immunoglobulin-like V-type domains.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X00412; CAA25118.1; .
DR PIR; A02111; QRRBG.
DR InterPro; IPR007110; Ig-1like.
DR InterPro; IPR003599; Ig.
DR InterPro; IPR003006; Ig_MHC.
DR Pfam; PF00047; Ig; 5.
DR SMART; SM00409; Ig; 5.
DR PROSITE; PS50835; IG-LIKE; 3.

```

```

KW Immunoglobulin domain; Repeat; Transmembrane; Glycoprotein; Signal;
KW Polymorphism.
FT SIGNAL 1 18
FT CHAIN 19 773 POLYMERIC-IMMUNOGLOBULIN RECEPTOR.
FT CHAIN 19 615 SECRETORY COMPONENT.
FT DOMAIN 19 647 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 648 670 POTENTIAL.
FT DOMAIN 671 773 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 25 131 IG-LIKE V-TYPE 1.
FT DOMAIN 138 232 IG-LIKE V-TYPE 2.
FT DOMAIN 233 340 IG-LIKE V-TYPE 3.
FT DOMAIN 352 455 IG-LIKE V-TYPE 4.
FT DOMAIN 461 557 IG-LIKE V-TYPE 5.
FT DISULFID 46 115 POTENTIAL.
FT DISULFID 155 225 POTENTIAL.
FT DISULFID 260 324 POTENTIAL.
FT DISULFID 369 438 POTENTIAL.
FT DISULFID 478 538 POTENTIAL.
FT CARBOHYD 88 88 N-LINKED (GLCNAc. . .) (IN ALLOTYPES T61)
FT CARBOHYD 108 108 N-LINKED (GLCNAc. . .) (IN ALLOTYPES T62
AND T63).
FT CARBOHYD 418 418 N-LINKED (GLCNAc. . .).
FT VARIANT 88 88 K -> N (IN ALLOTYPIC T61).
FT VARIANT 94 94 D -> E (IN ALLOTYPIC T61).
FT VARIANT 101 108 TVDQLTQN -> YLNRLSQS (IN ALLOTYPIC T61).
FT VARIANT 110 110 S -> T (IN ALLOTYPIC T63).
SQ SEQUENCE 773 AA; 83886 MW; DF2C44D2F1193C65 CRC64;

Query Match 1.7%; Score 13; DB 1; Length 773;
Best Local Similarity 100.0%; Pred. No. 9.5e-05;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 727 EPKAKRSSKEEA 739
    |||||
Db 736 EPKAKRSSKEEA 748

RESULT 6
PIGR_PIG STANDARD; PRT: 102 AA.
ID PIGR_PIG STANDARD; PRT: 102 AA.
AC Q29244;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Polymeric-immunoglobulin receptor (Poly-Ig receptor) (PIGR)
DE (Fragment).
GN PIGR.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A.
RX TISSUE=Small intestine;
RX MEDLINE=96327607; PubMed=8672129;
RA Wintrobe A.K., Fredholm M., Davies W.;
RT "Evaluation and characterization of a porcine small intestine cDNA
RT library: analysis of 839 clones.";
RL Mamm. Genome 7:509-517(1996).
CC -1- FUNCTION: THIS RECEPTOR BINDS POLYMERIC IGA AND IGM AT THE
CC BASOLATERAL SURFACE OF EPITHELIAL CELLS. THE COMPLEX IS THEN
CC TRANSPORTED ACROSS THE CELL TO BE SECRETED AT THE APICAL SURFACE.
CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN. ALSO SECRETED.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; F14851; CAA23294.1; .

```



Db 16 HIEGEPNL 23

RESULT 8  
HIS6\_SULSO STANDARD: PRT: 251 AA.  
AC O33774:  
DT 15-DEC-1998 (Rel. 37, Created)  
DT 15-DEC-1998 (Rel. 37, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Imidazole glycerol phosphate synthase subunit hisF (PC 4.1.3.-) (IGP synthase cyclase subunit) (IGP synthase subunit hisF) (ImGP synthase subunit hisF) (IGPS subunit hisF).  
DE HISF OR S500597 OR C08\_052.  
OS Sulfolobus solfataricus.  
OC Archaea: Crenarchaeota; Thermoprotei; Sulfolobales; Sulfolobaceae;  
OC Sulfolobus.  
OC NCBI-TaxID:2287;  
OX  
RN  
RN (1)  
RP SEQUENCE FROM N.A.  
RC STRAIN-ATCC 35092 / DSM 1617 / P2;  
RX MEDLINE-97352708; PubMed-920967;  
RA Charlebois R.L., Sensen C.W., Doolittle W.F., Brown J.R.;  
RT "Evolutionary analysis of the hisGABDFEHI gene cluster from the archaeon Sulfolobus solfataricus P2.";  
RL J. Bacteriol. 179:4429-4432(1997).  
RN (2)  
RP SEQUENCE FROM N.A.  
RC STRAIN-ATCC 35092 / DSM 1617 / P2;  
RX MEDLINE-20165948; PubMed-10701121;  
RA Charlebois R.L., Singh R.K., Chan-Weher C.C.-Y., Allard G., Chow C., Confalonieri F., Curtis B., Duguet M., Erasuo G., Faguy D., Gaasterland T., Garrett R.A., Gordon P., Jeffries A.C., Kozera C., Kishwaha N., Latteur E., Medina N., Peng X., Penny S.L., She Q., St Jean A., van der Oost J., Young F., Zivanovic Y., Doolittle W.F., Ragan M.A., Sensen C.W.;  
RT "Gene content and organization of a 281-kbp contig from the genome of the extremely thermophilic archaeon, Sulfolobus solfataricus P2.";  
RL Genome 43:116-136(2000).  
RN (3)  
RP SEQUENCE FROM N.A.  
RC STRAIN-ATCC 35092 / DSM 1617 / P2;  
RX MEDLINE-21332296; PubMed-11427726;  
RA She Q., Singh R.K., Confalonieri F., Zivanovic Y., Allard G., Aweyer M.J., Chan-Weher C.C.-Y., Clausen I.G., Curtis B.A., De Moors A., Erasuo G., Fletcher C., Gordon P.M.K., Heikamp-de Jong I., Jeffries A.C., Kozera C.J., Medina N., Peng X., Ra Thi-Ngoc H.P., Redder P., Schenk M.E., Theriault C., Tolstrup N., Charlebois R.L., Doolittle W.F., Duguet M., Gaasterland T., Garrett R.A., Ragan M.A., Sensen C.W., Van der Oost J.;  
RT "The complete genome of the crenarchaeon Sulfolobus solfataricus P2.";  
RL Proc. Natl. Acad. Sci. U.S.A. 98:7833-7840(2001).  
CC -1- FUNCTION: IGP synthase catalyzes the conversion of PRFAR and glutamine to IGP, AICAR and glutamate. The hisF subunit catalyzes the cyclization activity that produces IGP and AICAR from PRFAR using the ammonia provided by the hisH subunit (BY similarity).  
CC -1- CATALYTIC ACTIVITY: 5-[(5-phospho-1-deoxyribulose-1-ylamino)methyl]ideneamino-1'-1'-5-phosphoribosyl]imidazole-4-carboxamide + L-glutamine = imidazole-glycerol phosphate + 5-aminoimidazole-4-carboxamide ribonucleotide + L-glutamate + H(2O).  
CC -1- PATHWAY: Histidine biosynthesis; fifth step.  
CC -1- SUBUNIT: Heterodimer of hisH and hisF (BY similarity).  
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (BY similarity).  
CC -1- SIMILARITY: BELONGS TO THE HISA / HISF FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (see <http://www.isb-sdb.ch/announce/>  
CC or send an email to [license@sdb.ch](mailto:license@sdb.ch)).  
CC

DR EMBL: U82227; AAB63022.1; -  
 DR EMBL: Y18930; CAB57702.1; -  
 DR EMBL: AE006689; AAK40908.1; -  
 DR PIR: E90206; E90206.  
 DR HAMAP: MF\_01013; -; 1.  
 DR InterPro: IPR001009; FMN\_enzyme.  
 DR InterPro: IPR004651; HisF.  
 DR Pfam: PF00977; His\_biosynth; 1.  
 DR TIGRFAMs: TIGR00735; hisF; 1.  
 KW Histidine biosynthesis; Lyase; Complete proteome.  
 FT ACT\_SITE 11 11 POTENTIAL.  
 FT ACT\_SITE 130 130 POTENTIAL.  
 SQ SEQUENCE 251 AA; 26965 MW; 46246934AC1BA850 CRC64;

Query Match 1.0%; Score 8; DB 1; Length 251;  
 Best Local Similarity 100.0%; Pred. No. 5.4;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 574 SLAKADAA 581  
 DB 215 SLAKADAA 222

RESULT 9  
 T2M1\_MICAM STANDARD; PRT; 308 AA.  
 ID T2M1\_MICAM  
 AC P50189;  
 DT 01-OCT-1996 (Rel. 34, Created)  
 DT 01-OCT-1996 (Rel. 34, Last sequence update)  
 DT 15-SEP-2003 (Rel. 42, Last annotation update)  
 DE Type II restriction enzyme MamI (EC 3.1.21.4) (Ecdonuclease MamI)  
 DE (R.MamI).  
 GN MAMIR.  
 OS Microbacterium ammoniaphilum.  
 CC Bacteria; Actinobacteria; Actinomycetales;  
 CC Micrococcales; Microbacteriaceae; Microbacterium.  
 CC NCBI\_TaxID=42460;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=ATCC 15354;  
 RX MEDLINE=96257250; PubMed=8654988;  
 RA Striebel H.-W., Seeder S., Jarsch M., Kessler C.;  
 RT "Cloning and characterization of the MamI restriction-modification  
 RT system from Microbacterium ammoniaphilum in Escherichia coli.";  
 RL Gene 172:41-46(1996).

CC -1- FUNCTION: RECOGNIZES THE DOUBLE-STRANDED SEQUENCE GATNNNNATC AND  
 CC CLEAVES AFTER N-5.  
 CC -1- CATALYTIC ACTIVITY: Endonucleolytic cleavage of DNA to give  
 CC specific double-stranded fragments with terminal 5'-phosphates.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: X79027; CAAS5648.1; -  
 DR PIR: T45133; T45133.  
 DR REBASE: 1200; MamI.  
 KW Restriction system; Hydrolase; Nuclease; Endonuclease.  
 SO SEQUENCE 308 AA; 34293 MW; 53057207BFB1B92 CRC64;

Query Match 1.0%; Score 8; DB 1; Length 308;  
 Best Local Similarity 100.0%; Pred. No. 6.4;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 280 DVVVNTLG 287  
 DB 200 DVVVNTLG 207

RESULT 10  
 YD01\_METUA STANDARD; PRT; 400 AA.  
 ID YD01\_METUA  
 AC O58697;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Hypothetical ATP-binding protein MJ1301.  
 GN MJ1301.  
 OS Methanococcus jannaschii.  
 CC Archaea; Euryarchaeota; Methanococci; Methanococcales;  
 CC Methanocaldococcaceae; Methanocaldococcus.  
 CC NCBI\_TaxID=2190;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=JAL-1 / DSM 2661 / ATCC 43067;  
 RX MEDLINE=96337999; PubMed=8688087;  
 RA Buit C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,  
 RA Sutton G.G., Blake J.A., Fitzgerald L.M., Clayton R.A., Gocayne J.D.,  
 RA Kerlavage A.R., Dougherty B.A., Tomb J.-F., Adams M.D., Reich C.I.,  
 RA Overbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Glodok A.,  
 RA Scott J.L., Geoghagen N.S.M., Weidman J.F., Fuhrmann J.L., Nguyen D.,  
 RA Uterback T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,  
 RA Cotton M.D., Roberts K.M., Hurst M.A., Kalne B.P., Borodovsky M.,  
 RA Klenk H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.;  
 RT "Complete genome sequence of the methanogenic archaeon, Methanococcus  
 RT jannaschii.";  
 RL Science 273:1058-1073(1996).  
 RN [2]  
 RP SIMILARITY.  
 RC MEDLINE=97197912; PubMed=9045616;  
 RA Koornik E.V.;  
 RT "Evidence for a family of archaeal ATPases.";  
 RL Science 275:1489-1490(1997).

CC -1- SIMILARITY: BELONGS TO THE ARCHAEOAL ATPASE FAMILY.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: U67571; AAB99326.1; -  
 DR PIR: D64462; D64462.  
 DR TIGR: MJ1301; -  
 DR InterPro: IPR003593; AAA\_ATPase.  
 DR InterPro: IPR002576; Archaeal\_ATPase.  
 DR Pfam: PF01637; Archaeal\_ATPase; 1.  
 DR ProDom: PD003808; Archaeal\_ATPase; 1.  
 DR SMART: SM00382; AAA; 1.  
 KW Hypothetical protein; ATP-binding; Complete proteome.  
 FT NP\_BIND 36 43 ATP (POTENTIAL).  
 FT SEQUENCE 400 AA; 46903 MW; EE9460CAD791D049 CRC64;

Query Match 1.0%; Score 8; DB 1; Length 400;  
 Best Local Similarity 100.0%; Pred. No. 7.9;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 457 IIEGPNL 464  
 DB 23 IIEGPNL 30

RESULT 11  
 SYFB\_PYPAB STANDARD; PRT; 556 AA.  
 ID SYFB\_PYPAB  
 AC Q9UYX2;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 15-SEP-2003 (Rel. 42, Last annotation update)

```

DE Phenylalanyl-tRNA synthetase beta chain (EC 6.1.1.20) (Phenylalanine--
DE tRNA ligase beta chain) (PHERS).
GN PHE1 OR PYRAB13850 OR PAB2427.
OS PYROCoccus abyssi.
OC Archaea: Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;
CC Pyrococcus.
OX NCBI_TaxID=29292;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=GES / Orsay;
RX PubMed=12622808;
RA Cohen G.N., Barbe V., Flament D., Galperin M., Hellig R., Lecompte O.,
RA Poch O., Pileur D., Querellou J., Ripp R., Thierry J.-C.,
RA van der Oost J., Weissenbach J., Zivanovic Y., Forterre P.;
RT "An integrated analysis of the genome of the hyperthermophilic
RT archaeon Pyrococcus abyssi."
RL Mol. Microbiol. 47:1495-1512(2003).
CC -1- CATALYTIC ACTIVITY: ATP + L-phenylalanine + tRNA(Phe) = AMP +
CC diphosphate + L-phenylalanyl-tRNA(Phe).
CC -1- SUBUNIT: Tetramer of two alpha and two beta chains (By
CC similarity).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- SIMILARITY: BELONGS TO THE PHENYLALANYL-TRNA SYNTHETASE BETA
CC CHAIN FAMILY. SUBFAMILY 2.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL: AJ248287; CAB50290.1; -
DR PIR: E75049; E75049.
DR HAMAP: MF_00284; B3_1.
DR InterPro: IPR005147; B5.
DR InterPro: IPR004531; PheT_arch.
DR DR PF03484; B5; 1.
DR TIGRPFAMs: TIGR00471; PheT_arch; 1.
DR TIGRPFAMs: TIGR00471; PheT_arch; 1.
KW Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding;
KM Complete proteome.
FT DOMAIN 135 POLY-ARG.
FT SEQUENCE 556 AA; 63949 MW; 06134413484E60DB CRC64;
SQ
Query Match 1.0%; Score 8; DB 1; Length 556;
Best Local Similarity 100.0%; Pred. No. 10;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 396 PLYVDSSEG 403
Db 200 PLYVDSSEG 207
-----
RESULT 12
ID SYFB_PYREFU STANDARD; PRT; 556 AA.
AC O80260;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Phenylalanyl-tRNA synthetase beta chain (EC 6.1.1.20)
DE (Phenylalanine--tRNA ligase beta chain) (PHERS).
GN PHE1 OR PF0990.
OS Pyrococcus furiosus.
OC Archaea: Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;
CC Pyrococcus.
OX NCBI_TaxID=2261;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Vcl / DSM 3638 / ATCC 43587 / JCM 8422;
RA Weiss R.B., Dunn D.M., Robb F.T., Brown J.R.;
RT "The complete sequence of the Pyrococcus furiosus genome."

```

```

RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
CC -1- CATALYTIC ACTIVITY: ATP + L-phenylalanine + tRNA(Phe) = AMP +
CC diphosphate + L-phenylalanyl-tRNA(Phe).
CC -1- SUBUNIT: Tetramer of two alpha and two beta chains (By
CC similarity).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- SIMILARITY: BELONGS TO THE PHENYLALANYL-TRNA SYNTHETASE BETA CHAIN
CC FAMILY. SUBFAMILY 2.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL: AE010211; AL081114.1; -
DR HAMAP: MF_00284; B3_4.
DR InterPro: IPR005146; B3_4.
DR InterPro: IPR005147; B5.
DR InterPro: IPR004531; PheT_arch.
DR PF03483; B3_4; 1.
DR PF03484; B5; 1.
DR TIGRPFAMs: TIGR00471; PheT_arch; 1.
DR TIGRPFAMs: TIGR00471; PheT_arch; 1.
KW Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding;
KM Complete proteome.
FT SEQUENCE 556 AA; 63963 MW; CFFB32F6D001432 CRC64;
SQ
Query Match 1.0%; Score 8; DB 1; Length 556;
Best Local Similarity 100.0%; Pred. No. 10;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 396 PLYVDSSEG 403
Db 200 PLYVDSSEG 207
-----
RESULT 13
ID B2NB_PSEFL STANDARD; PRT; 563 AA.
AC P51853;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Benzaldehyde lyase (EC 4.1.2.38) (Benzoin aldolase) (BZL).
GN B2NB.
OS Pseudomonas fluorescens.
OC Bacteria: Proteobacteria; Gammaproteobacteria; Pseudomonadales;
CC Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxID=294;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Biovar I;
RX MEDLINE=94299157; PubMed=8026749;
RA Hinrichsen P., Gomez I., Vicuna R.;
RT "Cloning and sequencing of the gene encoding benzaldehyde lyase from
RT Pseudomonas fluorescens biovar I."
RL Gene 144:137-138(1994).
CC -1- FUNCTION: CLEAVAGE OF BENZOIN-ANISOIN LINKAGE.
CC -1- CATALYTIC ACTIVITY: Benzoin -> 2 benzaldehyde.
CC -1- COFACTOR: Binds 1 thiamine pyrophosphate and 1 metal ion per
CC subunit.
CC -1- SIMILARITY: BELONGS TO THE TPP ENZYMES FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----

```

DR EMBL: U04048; AAA50176.1; InterPro: IPR00399; Pyruvate, decarb.  
 DR Pfam: PF00205; TPP\_enzymes\_1.  
 DR Pfam: PF02775; TPP\_enzymes\_C; 1.  
 DR Pfam: PF02776; TPP\_enzymes\_N; 1.  
 KW Lyase: Flavoprotein: Thiamine pyrophosphate.  
 FT ACT\_SITE 50 50  
 FT SEQUENCE 563 AA: 58957 MW: 36940C0B33265DBF CRC64:  
 SQ SEQUENCE 563 AA: 58957 MW: 36940C0B33265DBF CRC64:  
 Query Match 1.0%; Score 8; DB 1; Length 563;  
 Best Local Similarity 100.0%; Pred. No. 11;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Oy 576 AKADAPP 583  
 Db 264 AKADAPP 271  
 RESULT 14  
 FGR4\_HUMAN STANDARD: PRT: 802 AA.  
 ID FGR4\_HUMAN  
 AC P22455; O43785; Q14309;  
 DT 01-AUG-1991 (Rel. 19, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 15-SEP-2003 (Rel. 42, Last annotation update)  
 DE Fibroblast growth factor receptor 4 precursor (EC 2.7.1.112) (FGFR-4).  
 GN FGFR4 OR JTK2 OR TKF.  
 OS Homo sapiens (Human).  
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 OX NCBI\_TaxId:9606;  
 RN (1)  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE-91224085; PubMed-1709094;  
 RA Paranen J.M., Mäkelä T.P., Berola E., Korhonen J., Hirvonen H.,  
 Claesson-Welsh L., Allitalo K.,  
 "FGR-4, a novel acidic fibroblast growth factor receptor with a  
 distinct expression pattern."  
 RT EMBO J. 10:1347-1354(1991).  
 RL (2)  
 RP SEQUENCE FROM N.A.  
 RC TISSUE-Mammary gland;  
 RX MEDLINE-93194827; PubMed-7680645;  
 RA Ron D., Reich R., Chedid M., Lengel C., Cohen O.E., Chan A.M.,  
 Neufeld G., Miki T., Tronick S.R.,  
 "Fibroblast growth factor receptor 4 is a high affinity receptor for  
 both acidic and basic fibroblast growth factor but not for  
 keratinocyte growth factor."  
 RT J. Biol. Chem. 268:5388-5394(1993).  
 RL (3)  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE-98119018; PubMed-9457674;  
 RA Kostzawa M., Müller U.,  
 "Genomic structure and complete sequence of the human FGFR4 gene."  
 RT Mamm. Genome 9:131-135(1998).  
 RL (4)  
 RP SEQUENCE FROM N.A.  
 RC TISSUE-Muscle;  
 RX MEDLINE-22388257; PubMed-12477932;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
 Altschul S.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
 Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
 Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 Brownstein M.J., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
 Rana S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,  
 Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Huliyk S.W.,  
 Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 Fahy J., Helton E., Kettner M., Madan A., Rodriguez S., Sanchez A.,  
 Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,

RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smallus D.E.,  
 RA Schermer A., Schein J.E., Jones S.J.M., Marra M.A.,  
 "Generation and initial analysis of more than 15,000 full-length  
 human and mouse cDNA sequences".  
 RT Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
 RN (5)  
 RP SEQUENCE OF 609-676 FROM N.A.  
 RC TISSUE-Blood;  
 RX MEDLINE-91062389; PubMed-2247464;  
 RA Paranen J., Mäkelä T.P., Allitalo R., Lehto-Salho H., Allitalo K.,  
 "Putative tyrosine kinases expressed in K-562 human leukemia cells."  
 RT Proc. Natl. Acad. Sci. U.S.A. 87:8913-8917(1990).  
 RL (1)  
 CC FUNCTION: RECEPTOR FOR ACIDIC FIBROBLAST GROWTH FACTOR. DOES NOT  
 CC BIND TO BASIC FIBROBLAST GROWTH FACTOR. BINDS FGF19.  
 CC -1- CATALYTIC ACTIVITY: ATP + a protein tyrosine -> ADP + protein  
 CC tyrosine phosphate.  
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.  
 CC -1- SIMILARITY: BELONGS TO THE FIBROBLAST GROWTH FACTOR RECEPTOR  
 CC FAMILY.  
 CC -1- SIMILARITY: Contains 3 immunoglobulin-like C2-type domains.  
 CC -----  
 CC EMBL: X57205; CAA0490.1; -  
 CC EMBL: L03840; AAB59389.1; -  
 CC EMBL: Y13901; CAB74200.1; -  
 CC EMBL: BC011847; AAH11847.1; -  
 CC EMBL: M59373; AA63208.1; -  
 CC PIR: S15345; TVHOF4.  
 CC PDB: 1OCT; 15-JUN-99.  
 CC Genew: HGNC:3691; FGFR4.  
 DR MIM: 134935; -  
 DR GO: GO:0005887; C: Integral to plasma membrane; TAS.  
 DR GO: GO:0005007; F: fibroblast growth factor receptor activity; TAS.  
 DR GO: GO:0008343; P: FGF receptor signaling pathway; TAS.  
 DR InterPro: IPR007110; 1g-1like.  
 DR InterPro: IPR003598; 1g-c2.  
 DR InterPro: IPR003006; 1g-MHC.  
 DR InterPro: IPR000719; Prot\_kinase.  
 DR InterPro: IPR001245; Tyr\_kinase.  
 DR Pfam: PF00047; 1g; 3.  
 DR Pfam: PF00069; PKinase; 1.  
 DR PRINTS: PR00109; TYRKINASE.  
 DR ProDom: PD000001; Prot\_kinase; 1.  
 DR SMART: SM00408; Igc2; 3.  
 DR SMART: SM00219; Tyrc; 1.  
 DR PROSITE: PS00835; 1g-LIKE; 2.  
 DR PROSITE: PS00107; PROTEIN\_KINASE\_ATP; 1.  
 DR PROSITE: PS50011; PROTEIN\_KINASE\_DOM; 1.  
 DR PROSITE: PS00109; PROTEIN\_KINASE\_TYR; 1.  
 KW Receptor: Glycoprotein: Tyrosine-protein kinase: ATP-binding;  
 KW transferase: Phosphorylation; Transmembrane; Immunoglobulin domain;  
 KW Repeat: Signal; Polymorphism; 3D-structure.  
 FT SIGNAL 1 24  
 FT CHAIN 25 802  
 FT DOMAIN 25 369  
 FT TRANSMEM 370 390  
 FT DOMAIN 391 802  
 FT DOMAIN 30 128  
 FT DOMAIN 152 240  
 FT DOMAIN 249 349  
 FT DOMAIN 467 755  
 FT NP\_BIND 473 481  
 FT BINDING 503 503  
 FT ACT\_SITE 612 612  
 FT MOD\_RES 643 643  
 P1: FIBROBLAST GROWTH FACTOR RECEPTOR 4.  
 P2: EXTRACELLULAR (POTENTIAL).  
 P3: POTENTIAL.  
 P4: CYTOPLASMIC (POTENTIAL).  
 P5: 1G-LIKE C2-TYPE 1.  
 P6: 1G-LIKE C2-TYPE 2.  
 P7: 1G-LIKE C2-TYPE 3.  
 P8: PROTEIN KINASE.  
 P9: ATP (BY SIMILARITY).  
 P10: ATP (BY SIMILARITY).  
 P11: BY SIMILARITY.  
 P12: PHOSPHORYLATION (AUTO-) (BY SIMILARITY).

```

FT DISULFID 57 101 POTENTIAL.
FT DISULFID 172 224 POTENTIAL.
FT DISULFID 271 333 POTENTIAL.
FT CARBOHYD 112 112 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 258 258 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 290 290 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 311 311 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 322 322 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT VARIANT 388 388 G -> R (IN OBSNP:351855).
FT /FTID-VAR.014797.
FT CONFLICT 297 297 D -> V (IN REF. 1).
SQ SEQUENCE 802 AA: 87954 MW: B2B259831BB889F CRC64:

Query Match
Best Local Similarity 100.0%; Score 8; DB 1; Length 802;
Pred. No. 14;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 318 EDAGRYLC 325
DB 94 EDAGRYLC 101

RESULT 15
FGF4_MOUSE STANDARD: PRT; 808 AA.
ID FGF4_MOUSE
AC 003142;
DT 01-OCT-1994 (Rel. 30, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Fibroblast growth factor receptor 4 precursor (EC 2.7.1.112) (FGFR-4)
DE (Protein-tyrosine kinase receptor MPK-11).
GN FGFR4 OR FGFR-4 OR MPK-11.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CD-1; TISSUE=Fetal cerebellum;
RX MEDLINE=92146274; Pubmed=1723680;
RA Stark K.L., McMahon J., McMahon A.P.;
RT "FGFR-4, a new member of the fibroblast growth factor receptor
RT family, expressed in the definitive endoderm and skeletal muscle
RT lineages of the mouse."
RL Development 113:641-651(1991).
RN [2]
RP SEQUENCE OF 620-676 FROM N.A.
RC STRAIN=C57BL/6; TISSUE=Embryonic brain;
RX MEDLINE=93096464; Pubmed=1281307;
RA Gilardi-Hebenstreit P., Nieto M.A., Fraaij M., Mattei M.-G.,
RA Cheslier A., Wilkinson D.G., Charney P.;
RT "An Eph-related receptor protein tyrosine kinase gene segmentally
RT expressed in the developing mouse hindbrain."
RL Oncogene 7:2499-2506(1992).
CC -1- FUNCTION: PUTATIVE RECEPTOR FOR BASIC FIBROBLAST GROWTH FACTOR.
CC MAY BE INVOLVED IN THE DEVELOPMENT OF SKELETAL MUSCLE CELL
CC LINEAGES.
CC -1- CATALYTIC ACTIVITY: ATP + a protein tyrosine - ADP + protein
CC tyrosine phosphate.
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC -1- TISSUE SPECIFICITY: EXPRESSED IN THE DEVELOPING GUT ENDODERM,
CC IN MYOTOMALLY DERIVED SKELETAL MUSCLE, THE ADRENAL CORTEX,
CC KIDNEY AND CONDENSING CARTILAGE. PRESENT IN ADULT LIVER,
CC LUNG AND KIDNEY.
CC -1- SIMILARITY: BELONGS TO THE FIBROBLAST GROWTH FACTOR RECEPTOR
CC FAMILY.
CC -1- SIMILARITY: Contains 3 immunoglobulin-like C2-type domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial

```

```

CC entitles requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: X59927; CAA42551.1; -
CC EMBL: X57236; CAA40512.1; -
CC F01; S18209; S18209.
CC HSSP: P11362; IFGK.
CC MGD: MGI:95525; Fgf4.
CC InterPro: IPR007110; Ig_Like.
CC InterPro: IPR003598; Ig_C2.
CC InterPro: IPR003006; Ig_MHC.
CC InterPro: IPR007119; Prot_Kinase.
CC InterPro: IPR01245; Tyr_PKinase.
CC Pfam: PF00047; Ig_3.
CC Pfam: PF00069; pKinase; 1.
CC PRINTS: PR00109; TYRKINASE.
CC ProDom: PD000001; Prot_kinase; 1.
CC SMART: SM00408; Jc2; 3.
CC SMART: SM00219; Tyrc; 1.
CC PROSITE: PS50835; IG_LIKE; 2.
CC PROSITE: PS00107; PROTEIN_KINASE_ATP; FALSE_NEG.
CC PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.
CC PROSITE: PS00109; PROTEIN_KINASE_TYR; 1.
CC Receptor: Glycoprotein; Tyrosine-protein kinase; ATP-binding;
CC Transferrase; Phosphorylation; Transmembrane; Immunoglobulin domain;
CC Repeat; Signal.
CC FT SIGNAL 1 18 POTENTIAL.
CC FT CHAIN 19 808 FIBROBLAST GROWTH FACTOR RECEPTOR 4.
CC FT DOMAIN 19 366 EXTRACELLULAR (POTENTIAL).
CC FT TRANSMEM 367 387 POTENTIAL.
CC FT DOMAIN 388 808 CYTOPLASMIC (POTENTIAL).
CC FT DOMAIN 53 138 IG-LIKE C2-TYPE 1.
CC FT DOMAIN 149 237 IG-LIKE C2-TYPE 2.
CC FT DOMAIN 246 346 IG-LIKE C2-TYPE 3.
CC FT DOMAIN 464 761 PROTEIN KINASE.
CC FT NP_BIND 470 478 ATP (BY SIMILARITY).
CC FT BINDING 509 509 ATP (BY SIMILARITY).
CC FT ACT_SITE 618 618 BY SIMILARITY.
CC FT MOD_RES 649 649 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
CC FT DISULFID 53 98 POTENTIAL.
CC FT DISULFID 169 221 POTENTIAL.
CC FT DISULFID 268 330 POTENTIAL.
CC FT CARBOHYD 109 109 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC FT CARBOHYD 255 255 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC FT CARBOHYD 287 287 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC FT CARBOHYD 308 308 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC FT CARBOHYD 319 319 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 808 AA: 89775 MW: E3F30B5E334EB6 CRC64:

```

Query Match 1.0%; Score 8; DB 1; Length 808;  
Best Local Similarity 100.0%; Pred. No. 14;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 318 EDAGRYLC 325  
DB 91 EDAGRYLC 98

Search completed: September 9, 2003, 23:17:47  
Job time : 29 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OK protein - protein search, using sw model

Run on: September 9, 2003, 23:14:42 ; Search time 113 Seconds

(without alignments)  
1744.710 Million cell updates/sec

Title: US-09-818-247-1

Perfect score: 764

Sequence: 1 MLFVLVLCCLAVPAISTKTS.....DFLLQSTVAEAADGPOEA 764

Scoring table:  
Gapop 60.0 , Gapext 60.0

Searched: 830525 seqs, 258052604 residues

Word size: 0

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database : SPTREMBL\_23:\*

1: sp\_archaea:\*

2: sp\_bacteria:\*

3: sp\_fungi:\*

4: sp\_human:\*

5: sp\_invertebrate:\*

6: sp\_mammal:\*

7: sp\_mhc:\*

8: sp\_organelle:\*

9: sp\_phage:\*

10: sp\_plant:\*

11: sp\_prodent:\*

12: sp\_virus:\*

13: sp\_vertebrate:\*

14: sp\_unclassified:\*

15: sp\_virus:\*

16: sp\_bacteriophage:\*

17: sp\_archaeal:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	663	86.8	764	4	081ZV7
2	30	3.9	39	4	081ZV7
3	22	2.9	360	6	08WJ21
4	17	2.2	758	6	09N2H7
5	10	1.3	82	6	08SQ82
6	10	1.3	217	11	08R544
7	10	1.3	308	17	09HPL3
8	10	1.2	362	12	08JX22
9	9	1.2	362	12	08JX17
10	9	1.2	362	12	08JX17
11	9	1.2	731	6	08SP16
12	9	1.2	733	6	08SQ83
13	8	1.0	48	16	09X7N4
14	8	1.0	122	16	092K99
15	8	1.0	140	6	08S0D8
16	8	1.0	155	16	09RF31

17	8	1.0	168	11	08C0K6	08C0K6 mus musculus
18	8	1.0	197	17	057983	057983 pyrococcus
19	8	1.0	283	2	09F212	09F212 streptococc
20	8	1.0	287	11	09D7B8	09D7B8 mus musculus
21	8	1.0	287	12	066659	066659 equine herp
22	8	1.0	287	16	09KUN5	09KUN5 vibrio chol
23	8	1.0	311	16	08XUL4	08XUL4 raietonia s
24	8	1.0	496	4	08N311	08N311 homo sapien
25	8	1.0	501	2	093NF2	093NF2 neisseria g
26	8	1.0	509	17	08PTG6	08PTG6 methanosarc
27	8	1.0	562	11	08R113	08R113 mus musculus
28	8	1.0	563	2	09F413	09F413 pseudomonas
29	8	1.0	565	10	09FHY8	09FHY8 arabidopsis
30	8	1.0	570	11	09DBB9	09DBB9 mus musculus
31	8	1.0	648	5	09V6J4	09V6J4 drosophila
32	8	1.0	654	5	090863	090863 caenorhabdi
33	8	1.0	694	16	P96574	P96574 bacillus su
34	8	1.0	799	11	08C1B8	08C1B8 mus musculus
35	8	1.0	799	11	08C3V5	08C3V5 mus musculus
36	8	1.0	802	4	08TDA0	08TDA0 homo sapien
37	8	1.0	830	5	09BKT7	09BKT7 caenorhabdi
38	8	1.0	1092	11	091ZT0	091ZT0 rattus norv
39	8	1.0	1363	11	091ZT1	091ZT1 rattus norv
40	8	1.0	1553	5	095TR0	095TR0 drosophila
41	8	1.0	2181	5	09VRA6	09VRA6 drosophila
42	7	0.9	36	6	09TPP8	09TPP8 leontopithe
43	7	0.9	58	16	09BDT7	09BDT7 rhizobium 1
44	7	0.9	63	10	09XTP0	09XTP0 arabidopsis
45	7	0.9	67	16	08NLT3	08NLT3 corynebacte

## ALIGNMENTS

RESULT 1

ID 081ZV7 PRELIMINARY; PRT: 764 AA.

AC 081ZV7;

DT 01-MAR-2003 (TREMblrel. 23, Created)

DT 01-MAR-2003 (TREMblrel. 23, Last sequence update)

DT 01-MAR-2003 (TREMblrel. 23, Last annotation update)

DE Hepatocellular carcinoma associated protein T56.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

OX NCBI\_TaxID=9606;

FN [1]

RP SEQUENCE FROM N.A.

RA Dong X., Pang X., Cheng W.;

RT "Cloning and characterization of hepatocellular carcinoma associated-

RT genes.";

RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.

DR EMBL: AF272149; AAN65630.1; -

SO SEQUENCE 764 AA; 83283 MW; 927461F8EB305C7 CRC64;

Query Match	86.8%; Score 663; DB 4; Length 764;
Best Local Similarity	99.9%; Pred. No. 0;
Matches	763; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
OY	1 MLFVLVLCCLAVPAISTKSPICGPEEVNSVEGNSVITCYPTSVNRHTRKWCQGA 60
DB	1 MLFVLVLCCLAVPAISTKSPICGPEEVNSVEGNSVITCYPTSVNRHTRKWCQGA 60
OY	61 RGGCITLISSEGVSSKAGRANLTNPENCTFVNTAQLSDDSGRKCGLSNGLS 120
DB	61 RGGCITLISSEGVSSKAGRANLTNPENCTFVNTAQLSDDSGRKCGLSNGLS 120
OY	121 FDSVLEVSOGPGLNDKRYTVTDGRTVTINCPRKTENAKRSLYQIGLYPVLVIDSS 180
DB	121 FDSVLEVSOGPGLNDKRYTVTDGRTVTINCPRKTENAKRSLYQIGLYPVLVIDSS 180
OY	181 GYVNPNTGRRLDIOGTGOLFESVYNOLRLSDAGGYLCAGDGSNNKNADLOVLPK 240
DB	181 GYVNPNTGRRLDIOGTGOLFESVYNOLRLSDAGGYLCAGDGSNNKNADLOVLPK 240

```

Db      181  GYVNPYRTGRIRIDIGTCQQLFSSVYINQLRLSDAGQYLQCAQDDSNKKNADLOVLRP 240
Qy      241  EPELYVEDLRGVSFTFHCALGPEVANYAKFLCRSSGENDVYVNTLGKRAPAFEGRIILLN 300
Db      241  EPELYVEDLRGVSFTFHCALGPEVANYAKFLCRSSGENDVYVNTLGKRAPAFEGRIILLN 300
Qy      301  PODKGSFSSVITTLGRLKEDAGRYLCGASHDGOLOESPLOAMOLFVNESTIPRSTPVK 360
Db      301  PODKGSFSSVITTLGRLKEDAGRYLCGASHDGOLOESPLOAMOLFVNESTIPRSTPVK 360
Qy      361  GYAGSSVAVLCPYNRESKSIKWCMEGAONGRCPLVDSEGMVAOYEGRLSLEEG 420
Db      361  GYAGSSVAVLCPYNRESKSIKWCMEGAONGRCPLVDSEGMVAOYEGRLSLEEG 420
Qy      421  NGFTFYILNQLTSRDAGFYWCLTNGDTLWRTVEIRIIEGEPNLKYPGNVAVLGETLVK 480
Db      421  NGFTFYILNQLTSRDAGFYWCLTNGDTLWRTVEIRIIEGEPNLKYPGNVAVLGETLVK 480
Qy      481  PCHFPCKSSSEYKWKCKMNNNGCALPSODEGPKAFVNCDENSRLVSLTLNLTATDEG 540
Db      481  PCHFPCKSSSEYKWKCKMNNNGCALPSODEGPKAFVNCDENSRLVSLTLNLTATDEG 540
Qy      541  WYWCYKQGHFGETAAVVAVERKAAGSRDYSLAKADAPEKVLDSGFREIENKAIO 600
Db      541  WYWCYKQGHFGETAAVVAVERKAAGSRDYSLAKADAPEKVLDSGFREIENKAIO 600
Qy      601  DPLFAEERKAVADTRDQADGSRASVDSGSEEGSSRALVSTLVPLGLAVAGAVAVG 660
Db      601  DPLFAEERKAVADTRDQADGSRASVDSGSEEGSSRALVSTLVPLGLAVAGAVAVG 660
Qy      661  ARARHRKNVDRVRSIRSYRTDISDFENSRFGANDNMGASSITQETSLGKKEEVAATTE 720
Db      661  ARARHRKNVDRVRSIRSYRTDISDFENSRFGANDNMGASSITQETSLGKKEEVAATTE 720
Qy      721  STTETKEPKKARSSKEEAEMAYKDFLLQSTTAAEAQODPOEA 764
Db      721  STTETKEPKKARSSKEEAEMAYKDFLLQSTTAAEAQODPOEA 764

RESULT 2
O8TE27
ID O8TE27 PRELIMINARY: PRT: 39 AA.
AC O8TE27;
DT 01-JUN-2002 (TREMBLrel. 21, Created)
DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Polymeric immunoglobulin receptor (Fragment).
GN PIGR.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=9735866;
RA Pliskurich J.F., Youngman K.R., Phillips K.M., Hempen P.M.,
RA Blanchard M.H., France J.A., Kaetzel C.S.;
RT "Transcriptional regulation of the human polymeric immunoglobulin
RT receptor gene by interferon-gamma.";
RL Mol. Immunol. 34:75-91(1997).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=21270451; PubMed=11376148;
RA Bruce S.R., Peterson M.L.;
RT "Multiple features contribute to efficient constitutive splicing of an
RT unusually large exon.";
RL Nucleic Acids Res. 29:2292-2302(2001).
DR EMBL: AF261083; AAF72193.1; -
KM Receptor.
FT NON_TER 1 1
FT NON_TER 39 39
SQ SEQUENCE 39 AA: 4272 MW: 3FCBE72AA6467810 CRC64:

```

```

Query Match 3.9%; Score 30; DB 4; Length 39;
Best Local Similarity 100.0%; Pred. No. 1.1e-22;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      319  DAGRYLCGASHDGOLOESPIQAMOLFVNE 348
Db      10  DAGRYLCGASHDGOLOESPIQAMOLFVNE 39

RESULT 3
O8MJ21
ID O8MJ21 PRELIMINARY: PRT: 360 AA.
AC O8MJ21;
DT 01-OCT-2002 (TREMBLrel. 22, Created)
DT 01-OCT-2002 (TREMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE Polymeric immunoglobulin receptor (Fragment).
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX NCBI_TaxID=9615;
RN [1]
RP SEQUENCE FROM N.A.
RA Peters I.R., Helps C.R., Hall E.J., Day M.J.;
RT "Cloning and sequencing of canine polymeric immunoglobulin receptor
RT (PIgR) mRNA.";
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL: AY081057; AAL91653.1; -
DR InterPro: IPR003599; Ig.
DR InterPro: IPR003006; Ig_MHC.
DR Pfam: PF00047; Ig_2.
DR SMART: SM00409; Ig_1.
KW Receptor.
FT NON_TER 1 1
FT NON_TER 360 360
SQ SEQUENCE 360 AA: 39685 MW: 32AE72C03FA935CA CRC64:

Query Match 2.9%; Score 22; DB 6; Length 360;
Best Local Similarity 100.0%; Pred. No. 1.5e-13;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      104  DSGRYKGLGINSRGLSPDVS 125
Db      45  DSGRYKGLGINSRGLSPDVS 66

RESULT 4
O9NZH7
ID O9NZH7 PRELIMINARY: PRT: 758 AA.
AC O9NZH7;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE Poly-Ig receptor precursor.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A.
RA Sone T., Kumura H.;
RT "Porcine mammary gland cDNA clone, similar to poly-Ig receptor.";
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: AB032195; BAA84283.2; -
DR InterPro: IPR003599; Ig.
DR InterPro: IPR007110; Ig_Like.
DR InterPro: IPR003006; Ig_MHC.
DR Pfam: PF00047; Ig_5.
DR SMART: SM00409; Ig_4.
DR PROSITE: PS50835; IG_LIKE; 2.
KM Receptor; Signal.
FT SIGNAL 1 18
FT SIGNAL 758 758
SQ SEQUENCE 758 AA: 83154 MW: D5BBE1A8B082D247 CRC64:

```

Query Match 2.2%; Score 17; DB 6; Length 758;  
 Best Local Similarity 100.0%; Pred. No. 4.4e-08;  
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 116 SRGLSFDVSLVSGPG 132  
 DB 116 SRGLSFDVSLVSGPG 132

## RESULT 5

O8S082 PRELIMINARY: PRT; 82 AA.  
 AC O8S082;  
 DT 01-JUN-2002 (TREMBlrel. 21, Created)  
 DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)  
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)  
 DE Polymeric immunoglobulin receptor (Fragment).  
 GN PIGR.  
 OS Ovis aries (Sheep).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
 OC Bovidae; Caprinae; Ovis.  
 OX NCBI\_TaxID=9940;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Pfaffli M.W., Dzidic A., Rojas P.G., Bruckmaier R.M., Schams D.;  
 RT "Effects of an induced mammarygenesis and lactogenesis in sheep on the  
 RT mRNA expression levels of immune globulin receptors (FcRn, PIGR) and  
 RT zeta occludens proteins (Occludin, ZO1, ZO2, ZO3).";  
 RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AJ313189; CAC41993.1;  
 KW Receptor.  
 FT NON\_TER  
 FT NON\_TER  
 SQ SEQUENCE 82 AA; 8846 MW; D6040018954A50AA CRC64;

Query Match 1.3%; Score 10; DB 6; Length 82;  
 Best Local Similarity 100.0%; Pred. No. 0.11;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 145 GRTVINCIPF 154  
 DB 23 GRTVINCIPF 32

## RESULT 6

O8R544 PRELIMINARY: PRT; 217 AA.  
 AC O8R544;  
 DT 01-JUN-2002 (TREMBlrel. 21, Created)  
 DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)  
 DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)  
 DE Polymeric immunoglobulin receptor (Fragment).  
 GN PIGR.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA STRAIN-129/SV;  
 RC MEDLINE-99377172;  
 RX Bruce S.R., Kaetzel C.S., Peterson M.L.;  
 RT "Cryptic intron activation within the large exon of the mouse  
 RT polymeric immunoglobulin receptor gene: cryptic splice sites  
 RT correspond to protein domain boundaries.";  
 RL Nucleic Acids Res. 27:3446-3454(1999).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-129/SV;  
 RX MEDLINE-21270451; PubMed-11376148;  
 RA Bruce S.R., Peterson M.L.;

RT "Multiple features contribute to efficient constitutive splicing of an  
 RT unusually large exon.";  
 RT Nucleic Acids Res. 29:2292-2302(2001).  
 DR EMBL: AF261084; AAF72194.1;  
 DR MGD: MGI:103080; PIGR.  
 DR InterPro: IPR003599; Ig.  
 DR SMART: SM00409; IG; 2.  
 KW Receptor.  
 FT NON\_TER  
 FT NON\_TER  
 SQ SEQUENCE 217 AA; 24352 MW; B579C870ACFF180FB CRC64;

Query Match 1.3%; Score 10; DB 11; Length 217;  
 Best Local Similarity 100.0%; Pred. No. 0.25;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 312 ITGLKREDAG 321  
 DB 181 ITGLKREDAG 190

## RESULT 7

O9HPL3 PRELIMINARY: PRT; 308 AA.  
 ID O9HPL3;  
 AC O9HPL3;  
 DT 01-MAR-2001 (TREMBlrel. 16, Created)  
 DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)  
 DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)  
 DE Vng1578h.  
 GN VNG1578H.  
 OS Halobacterium sp. (strain NRC-1).  
 OC Archaea; Euryarchaeota; Halobacteria; Halobacteriales;  
 OC Halobacteriaceae; Halobacterium.  
 OX NCBI\_TaxID=64091;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA MEDLINE-20504483; PubMed-11016950;  
 RA Ng W.V., Kennedy S.P., Mahairas G.G., Bergquist B., Pan M.,  
 RA Shukla H.D., Lasry S.R., Baliga N.S., Thorsson V., Shroga J.,  
 RA Swartzell S., Weir D., Hall J., Dahl T.A., Weltl R., Goo Y.A.,  
 RA Lettner B., Keller K., Cruz R., Danson M.J., Hough D.W.,  
 RA Maddocks D.G., Jablonski P.E., Krebs M.P., Angewine C.M., Dale H.,  
 RA Isenbarger T.A., Peck R.F., Pohlischer M., Spudis J.L., Jung K.-H.,  
 RA Alam M., Freitas T., Hou S., Daniels C.J., Dennis P.P., Omer A.D.,  
 RA Ehardt H., Lowe T.M., Liang P., Riley M., Hood L., Dassarma S.;  
 RT "Genome sequence of Halobacterium species NRC-1.";  
 RT Proc. Natl. Acad. Sci. U.S.A. 97:12176-12181(2000).  
 DR EMBL: AE005070; AAG19854.1;  
 DR InterPro: IPR004485; CblB.  
 DR Pfam: PF03186; CblB\_CblB; 1.  
 KW Complete proteome.  
 SQ SEQUENCE 308 AA; 30446 MW; 3FA8778BA72D7ECB CRC64;

Query Match 1.3%; Score 10; DB 17; Length 308;  
 Best Local Similarity 100.0%; Pred. No. 0.34;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 652 AVGAVAVGVA 661  
 DB 5 AVGAVAVGVA 14

## RESULT 8

O8JX22 PRELIMINARY: PRT; 362 AA.  
 ID O8JX22;  
 AC O8JX22;  
 DT 01-OCT-2002 (TREMBlrel. 22, Created)  
 DT 01-OCT-2002 (TREMBlrel. 22, Last sequence update)  
 DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)  
 DE Polyprotein (Fragment).  
 GN P1.  
 OS Turnip mosaic virus (strain Japanese) (TuMV).  
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Potyviridae;



OC Poliovirus.  
 OX NCBI\_TaxID-12230;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-AL;  
 RX MEDLINE-22024854; PubMed-12029167;  
 RA Ohshtima K., Yamaguchi Y., Hirota R., Hamamoto T., Tomimura K., Tan Z.,  
 RA Sano T., Azuhata F., Walsh J.A., Fletcher J., Chen J., Gera A.,  
 RA Gibbs A.;  
 RT "Molecular evolution of Turnip mosaic virus: evidence of host  
 adaptation, genetic recombination and geographical spread.";  
 RL J. Gen. Virol. 83:1511-1521(2002).  
 DR EMBL: AB076441; BAC02777.1;  
 DR InterPro: IPR002540; Poty\_P1.  
 DR Pfam: PF01577; Poty\_P1; 1.  
 FT NON\_TER 362  
 SQ SEQUENCE 362 AA; 40817 MW; FD2B9F108A1870E8 CRC64;

Query Match 1.2%; Score 9; DB 12; Length 362;  
 Best Local Similarity 100.0%; Pred. No. 4.3;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 358 VVKGVAGSS 366  
 Db 63 VVKGVAGSS 71

RESULT 9  
 O8JX17 PRELIMINARY; PRT; 362 AA.  
 AC O8JX17;  
 DT 01-OCT-2002 (TREMBlrel. 22, Created)  
 DT 01-OCT-2002 (TREMBlrel. 22, Last sequence update)  
 DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)  
 DE Polyprotein (Fragment).  
 GN P1.  
 OS Turnip mosaic virus (strain Japanese) (TUNV).  
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Potyviridae;  
 OC Potyvirus.  
 OX NCBI\_TaxID-12230;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-1S1;  
 RX MEDLINE-22024854; PubMed-12029167;  
 RA Ohshtima K., Yamaguchi Y., Hirota R., Hamamoto T., Tomimura K., Tan Z.,  
 RA Sano T., Azuhata F., Walsh J.A., Fletcher J., Chen J., Gera A.,  
 RA Gibbs A.;  
 RT "Molecular evolution of Turnip mosaic virus: evidence of host  
 adaptation, genetic recombination and geographical spread.";  
 RL J. Gen. Virol. 83:1511-1521(2002).  
 DR EMBL: AB076446; BAC02782.1;  
 DR InterPro: IPR002540; Poty\_P1.  
 DR Pfam: PF01577; Poty\_P1; 1.  
 FT NON\_TER 362  
 SQ SEQUENCE 362 AA; 40466 MW; ACE990B0B9CE8A44 CRC64;

Query Match 1.2%; Score 9; DB 12; Length 362;  
 Best Local Similarity 100.0%; Pred. No. 4.3;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 358 VVKGVAGSS 366  
 Db 63 VVKGVAGSS 71

RESULT 10  
 O8JWY1 PRELIMINARY; PRT; 362 AA.  
 AC O8JWY1;  
 DT 01-OCT-2002 (TREMBlrel. 22, Created)  
 DT 01-OCT-2002 (TREMBlrel. 22, Last sequence update)  
 DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)  
 DE Polyprotein (Fragment).

GN P1.  
 OS Turnip mosaic virus (strain Japanese) (TUNV).  
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Potyviridae;  
 OC Potyvirus.  
 OX NCBI\_TaxID-12230;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-GBR7;  
 RX MEDLINE-22024854; PubMed-12029167;  
 RA Ohshtima K., Yamaguchi Y., Hirota R., Hamamoto T., Tomimura K., Tan Z.,  
 RA Sano T., Azuhata F., Walsh J.A., Fletcher J., Chen J., Gera A.,  
 RA Gibbs A.;  
 RT "Molecular evolution of Turnip mosaic virus: evidence of host  
 adaptation, genetic recombination and geographical spread.";  
 RL J. Gen. Virol. 83:1511-1521(2002).  
 DR EMBL: AB076482; BAC02818.1;  
 DR InterPro: IPR002540; Poty\_P1.  
 DR Pfam: PF01577; Poty\_P1; 1.  
 FT NON\_TER 362  
 SQ SEQUENCE 362 AA; 40267 MW; B991D09CA4C85F25 CRC64;

Query Match 1.2%; Score 9; DB 12; Length 362;  
 Best Local Similarity 100.0%; Pred. No. 4.3;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 358 VVKGVAGSS 366  
 Db 63 VVKGVAGSS 71

RESULT 11  
 O8SPI6 PRELIMINARY; PRT; 731 AA.  
 AC O8SPI6;  
 DT 01-JUN-2002 (TREMBlrel. 21, Created)  
 DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)  
 DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)  
 DE Polymetric immunoglobulin receptor.  
 GN PIGR.  
 OS Macropus eugenii (Tamar wallaby).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Metatheria; Diprotodontia; Macropodidae; Macropus.  
 OX NCBI\_TaxID-9315;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE-Mammary lymph node;  
 RA Taylor C.L., Harrison G.A.;  
 RT "cDNA cloning of the polymetric immunoglobulin receptor of the  
 marsupial Macropus eugenii (Tamar wallaby).";  
 RL Submitted (OCT-2000) to the EMBL/Genbank/DBD databases.  
 DR EMBL: AF317205; AAK69593.1;  
 DR InterPro: IPR003599; Ig.  
 DR InterPro: IPR007110; Ig\_Like.  
 DR InterPro: IPR003006; Ig\_MHC.  
 DR Pfam: PF00047; Ig; 5.  
 DR SMART: SM00409; Ig; 5.  
 DR PROSITE: PSS0835; IG\_LIKE; 3.  
 KW Receptor.  
 SQ SEQUENCE 731 AA; 80253 MW; 91BDF3BC3478BC17 CRC64;

Query Match 1.2%; Score 9; DB 6; Length 731;  
 Best Local Similarity 100.0%; Pred. No. 8;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 424 FTVIINQLT 432  
 Db 420 FTVIINQLT 428

RESULT 12  
 O8SOB3 PRELIMINARY; PRT; 733 AA.  
 ID O8SOB3  
 AC O8SOB3;

DT 01-JUN-2002 (TREMblrel. 21, Created)  
 DT 01-JUN-2002 (TREMblrel. 21, Last sequence update)  
 DT 01-MAR-2003 (TREMblrel. 23, Last annotation update)  
 DE Polymer-C-Immunoglobulin receptor.  
 GN PIGR.  
 OS Trichosurus vulpecula (Brush-tailed possum).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Metatheria; Diprotodontia; Phalangeridae; Trichosurus.  
 NCBI\_TaxID=9337;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Mammary gland;  
 RA Adamski F.M., Demmer J.;  
 RT "Cloning and characterization of pigr and j chain of the marsupial,  
 RT Trichosurus vulpecula (brush-tailed possum).";  
 RL Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AF091137; AAD1688.1;  
 DR InterPro: IPR003599; IG\_1.  
 DR InterPro: IPR007110; IG\_1like.  
 DR InterPro: IPR003006; IG\_MHC.  
 DR Pfam: PF00047; IG\_5.  
 DR SMART: SM00409; IG\_5.  
 DR PROSITE: PS50835; IG\_LIKE; 3.  
 SK Receptor.  
 SQ SEQUENCE 733 AA: 80300 MW: 8168872DBAC14A5A CRC64:

Query Match 1.0%; Score 9; DB 6; Length 733;  
 Best Local Similarity 100.0%; Pred. No. 8;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 424 FTVILNQLT 432  
 DB 422 FTVILNQLT 430

RESULT 13

O9X7N4 PRELIMINARY; PRT; 48 AA.  
 AC O9X7N4;  
 DT 01-NOV-1999 (TREMblrel. 12, Created)  
 DT 01-NOV-1999 (TREMblrel. 12, Last sequence update)  
 DT 01-MAR-2003 (TREMblrel. 23, Last annotation update)  
 DE Small hypothetical protein.  
 GN SC6728 OR SC5F2A.11.  
 OS Streptomyces coelicolor.  
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;  
 OC Streptomycineae; Streptomycetaceae; Streptomycetes.  
 NCBI\_TaxID=1902;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=A3(2) / M145;  
 RX MEDLINE=21996410; PubMed=12000953;  
 RA Bertley S.D., Chater K.F., Cerdano-Tarraga A.-M., Challis G.L.,  
 RA Thomson N.R., James K.D., Harris D.E., Quail M.A., Kleser H.,  
 RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,  
 RA Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,  
 RA Huang C.-H., Kleser T., Larke L., Murphy L., Oliver K., O'Neill S.,  
 RA Rabinowitsch E., Rajandream M.A., Rutherford K., Rutter S.,  
 RA Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,  
 RA Warren T., Wietzorrek A., Woodward J., Barrell B.G., Parkhill J.,  
 RA Hopwood D.A.;  
 RT "Complete genome sequence of the model actinomycete Streptomyces  
 RT coelicolor A3(2).";  
 RL Nature 417:141-147(2002).  
 DR EMBL: AL939129; CAB40678.1;  
 KW Hypothetical protein; Complete proteome.  
 SQ SEQUENCE 48 AA: 4983 MW: 93710C77F7944475 CRC64;

Query Match 1.0%; Score 8; DB 16; Length 48;  
 Best Local Similarity 100.0%; Pred. No. 7;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 648 GLVLAVGA 655

DB 24 GLVLAVGA 31

RESULT 14

O92K99 PRELIMINARY; PRT; 122 AA.  
 AC O92K99;  
 DT 01-DEC-2001 (TREMblrel. 19, Created)  
 DT 01-DEC-2001 (TREMblrel. 19, Last sequence update)  
 DT 01-OCT-2002 (TREMblrel. 22, Last annotation update)  
 DE Hypothetical protein R01815.  
 GN R01815 OR SMC00191.  
 OS Rhizobium meliloti (Sinorhizobium meliloti).  
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;  
 OC Rhizobiaceae; Sinorhizobium.  
 NCBI\_TaxID=382;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=1021;  
 RX MEDLINE=21396507; PubMed=11481430;  
 RA Capela D., Barloy-Hubler F., Gouzy J., Bothe G., Ampe F., Batut J.,  
 RA Boistard P., Becker A., Boutry M., Cadieu E., Dreano S., Gloux S.,  
 RA Godrie T., Goffeau A., Kahn D., Kiss E., Lelaure V., Masuy D.,  
 RA Pohl T., Portetelle D., Puhler A., Purnelle B., Rampsperger U.,  
 RA Renard C., Thebaud P., Vandenbol M., Weidner S., Gallibert F.;  
 RT "Analysis of the chromosome sequence of the legume symbiont  
 RT Sinorhizobium meliloti strain 1021.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 98:9877-9882(2001).  
 DR EMBL: AL591788; CAC46394.1;  
 KW Hypothetical protein; Complete proteome.  
 SQ SEQUENCE 122 AA: 12774 MW: F3C77A35579D8C21 CRC64;

Query Match 1.0%; Score 8; DB 16; Length 122;  
 Best Local Similarity 100.0%; Pred. No. 18;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 651 LAVGAVAV 658  
 DB 79 LAVGAVAV 86

RESULT 15

O8S0D8 PRELIMINARY; PRT; 140 AA.  
 AC O8S0D8;  
 DT 01-JUN-2002 (TREMblrel. 21, Created)  
 DT 01-JUN-2002 (TREMblrel. 21, Last sequence update)  
 DT 01-MAR-2003 (TREMblrel. 23, Last annotation update)  
 DE Vascular endothelial growth factor receptor 3 (Fragment).  
 OS Oryctolagus cuniculus (Rabbit).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.  
 NCBI\_TaxID=9986;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Yoon Y.-S., Murayama T., Isner J.M.;  
 RT "Favorable effect of VEGF-C Gene Transfer on Lymphedema.";  
 RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AF453570; AAL78955.1;  
 DR InterPro: IPR007110; IG\_1like.  
 DR InterPro: IPR003598; IG\_C2.  
 DR InterPro: IPR003006; IG\_MHC.  
 DR Pfam: PF00047; IG\_1.  
 DR SMART: SM00408; IGC2; 1.  
 DR PROSITE: PS50835; IG\_LIKE; 1.  
 KW Immunoglobulin domain; Receptor.  
 FT NON\_TER 1  
 FT TER 1  
 SQ SEQUENCE 140 AA: 15340 MW: C95ADA0160981E11 CRC64;

Query Match 1.0%; Score 8; DB 6; Length 140;  
 Best Local Similarity 100.0%; Pred. No. 20;

Matches	8;	Conservative	0;	Mismatches	0;	Indels	0;	Gaps	0;
Qy	318	EDAGRYLC	325						
Db	44	EDAGRYLC	51						

Search completed: September 9, 2003, 23:20:39  
Job time : 116 secs

---

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 9, 2003, 23:06:57 ; Search time 84 Seconds

(without alignments)  
1443.656 Million cell updates/sec

Title: US-09-818-247-1

Perfect score: 764  
Sequence: 1 MLFLVTLCLLAVFPAISTKS.....DFLLQSTVAEAADGPOEA 764

Scoring table: OLIGO  
Gapop 60.0 , Gapext 60.0

Searched: 1107863 seqs, 158726573 residues

Word size : 0

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

A\_Geneseq\_19Jun03.\*  
1: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.\*  
2: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.\*  
3: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.\*  
4: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.\*  
5: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.\*  
6: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.\*  
7: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.\*  
8: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.\*  
9: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.\*  
10: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.\*  
11: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.\*  
12: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.\*  
13: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.\*  
14: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.\*  
15: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.\*  
16: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.\*  
17: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.\*  
18: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.\*  
19: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.\*  
20: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.\*  
21: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.\*  
22: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.\*  
23: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.\*  
24: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2003.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	764	100.0	764	22	AA65695
2	764	100.0	764	23	AA65711
3	764	100.0	764	23	ABJ04350
4	746	97.6	746	17	AA03178
5	607	79.5	607	20	AA095601
6	605	79.2	607	20	AAV34099
7	605	79.2	607	23	AA047867
8	254	33.2	272	20	AAV73981
9	243	31.8	243	24	ABP55307

10	243	31.8	243	24	ABP55315	Human polyimmunog1
11	94	12.3	94	24	ABP55311	Human polyimmunog1
12	90	11.8	90	22	AA65712	Human polymeric im
13	61	8.0	61	19	AA043098	Polymeric immunog1
14	61	8.0	61	19	AA043099	Polymeric immunog1
15	60	7.9	60	20	AA09073	Immunoglobulin 11k
16	55	7.2	243	23	AAE29187	Cynomolgus monkey
17	55	7.2	243	24	ABP55308	Simian polyimmunog
18	55	7.2	243	24	ABP55316	Simian polyimmunog
19	55	7.2	243	24	ABP55317	Simian polyimmunog
20	50	6.5	70	21	AA053678	Human colon cancer
21	44	5.8	243	24	ABP55306	Polyimmunoglobulin
22	43	5.6	43	20	AA09078	Immunoglobulin 11k
23	31	4.1	31	19	AA043095	Human polymeric im
24	31	4.1	31	22	AA65713	Human polymeric im
25	31	4.1	243	24	ABP55314	Polyimmunoglobulin
26	29	3.8	46	20	AA09077	Immunoglobulin 11k
27	27	3.5	57	20	AA09072	Immunoglobulin 11k
28	22	2.9	94	24	ABP55310	Polyimmunoglobulin
29	21	2.7	757	17	AA03179	Bovine poly-immuno
30	21	2.7	757	22	AA65696	Polyimmunoglobulin
31	20	2.6	94	24	ABP55312	Simian polyimmunog
32	19	2.5	19	20	AA093888	Bifidobacterium bi
33	19	2.5	40	19	AA043097	Rabbit polymeric 1
34	18	2.4	18	23	ABG60662	Polyimmunoglobulin
35	18	2.4	23	23	ABG60663	Polyimmunoglobulin
36	18	2.4	23	23	ABG60664	Polyimmunoglobulin
37	17	2.2	769	17	AA03181	Rat poly-immunoglo
38	17	2.2	769	22	AA65697	Rat polymeric immu
39	17	2.2	771	23	ABG60638	Chimeric polyimmun
40	16	2.1	16	19	AA064615	Human polyimmunog1
41	16	2.1	16	19	AA061592	Polyimmunoglobulin
42	16	2.1	16	20	AA085768	Polyimmunoglobulin
43	16	2.1	16	23	ABG94837	Human polyimmunog1
44	16	2.1	16	23	ABG68266	Targeting molecuol
45	16	2.1	771	17	AA03180	Mouse poly-immunog

#### ALIGNMENTS

RESULT 1	
AA65695	AA65695 standard; protein: 764 AA.
XX	XX
AC	AA65695;
XX	XX
DT	07-JAN-2002 (first entry)
XX	XX
DE	Human polymeric immunoglobulin receptor (p1gR) sequence.
XX	XX
KW	Polymeric immunoglobulin receptor; p1gR; ligand; therapeutic;
KW	Carcinoma diagnosis; veterinary; human.
XX	XX
OS	Homo sapiens.
XX	XX
FT	FT
Key	Location/Qualifiers
FT	Peptide 487..603
FT	/note- "peptide to which a ligand binds to (claim 8)"
FT	Peptide 487..607
FT	/note- "peptide to which a ligand binds to"
FT	Peptide 487..611
FT	/note- "peptide to which a ligand binds to"
FT	Peptide 487..615
FT	/note- "peptide to which a ligand binds to"
FT	Peptide 487..618
FT	/note- "peptide to which a ligand binds to"
FT	Peptide 520..607
FT	/note- "peptide to which a ligand binds to"
FT	Peptide 520..611
FT	/note- "peptide to which a ligand binds to"
FT	Peptide 520..615
FT	/note- "peptide to which a ligand binds to"



KW Polymetric immunoglobulin receptor; pigR; ligand; therapeutic;  
 KW carcinoma diagnosis; veterinary; human.  
 XX Homo sapiens.  
 XX WO200172846-A2.  
 PD 04-OCT-2001.  
 XX 26-MAR-2001; 2001WO-US09699.  
 XX 27-MAR-2000; 2000US-192197P.  
 PR 27-MAR-2000; 2000US-192198P.  
 XX (REGC ) UNIV CALIFORNIA.  
 PA Mostov KE, Chapin SJ, Richman-Eisenstat J;  
 PI WPI; 2001-611619/70.  
 XX New ligands binding to a specific region of a polymetric immunoglobulin  
 PT receptor, useful for transporting therapeutic or diagnostic  
 PT compositions into or across cells expressing pigR e.g. in drug delivery  
 XX  
 XX Disclosure; Fig 2; 102pp; English.  
 PS The invention provides ligands that bind specifically to a region of an  
 XX animal cell polymetric immunoglobulin receptor (pigR). The pigR cleaves  
 CC to produce a stalk region remaining attached to the cell and a secretory  
 CC component existing in the organ of interest in several forms. The ligands  
 CC do not bind to the stalk or the most abundant form of the secretory  
 CC component present in the organ under physiological conditions. The  
 CC ligands are useful for transporting therapeutic or diagnostic  
 CC compositions into or across cells expressing pigR, useful to introduce  
 CC or transport ligands such as antibodies and/or to deliver biologically  
 CC active components such as proteins, nucleic acids or detectable labels.  
 CC They are used to deliver therapeutic compositions to mucosal surfaces  
 CC such as the gastro-intestinal tract, respiratory system etc. in humans.  
 CC They are also useful to label cells expressing pigR, e.g. to distinguish  
 CC epithelial cells from a mixed cell population in pathology studies or to  
 CC aid in carcinoma diagnosis (since pigR expression is reduced in  
 CC carcinomas relative to normal epithelium). They can also be used to  
 CC deliver veterinary compositions, especially in mammals such as farm,  
 CC domestic or wild mammals or birds e.g. birds reared for human  
 CC consumption. The present sequence represents a human pigR sequence.  
 CC  
 XX  
 SO Sequence 764 AA:  
 Query Match 100.0%; Score 764; DB 22; Length 764;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 764; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 301 PODKGSFVSIVITGLRKEDAGRYLCAHSDGLOEBSPIOAMOLFVNEESTIRSPYVK 360  
 DB 301 PODKGSFVSIVITGLRKEDAGRYLCAHSDGLOEBSPIOAMOLFVNEESTIRSPYVK 360  
 QY 361 GVAGSSVAVLCPRNRRESKIKYKCLMEGONRCFLVDSEGVAAQVREGRLSLLEBP 420  
 DB 361 GVAGSSVAVLCPRNRRESKIKYKCLMEGONRCFLVDSEGVAAQVREGRLSLLEBP 420  
 QY 421 NGFTVILNQLTSDRDAFYWCLTNGDTLMTFTVEIKIIEGEPILKYPGNVAVLGETLKY 480  
 DB 421 NGFTVILNQLTSDRDAFYWCLTNGDTLMTFTVEIKIIEGEPILKYPGNVAVLGETLKY 480  
 QY 481 PCHEPCKFSSYEKWKWNNFGQALPSODEGSKAFVNCDENSRLVSLTLNVTBDEG 540  
 DB 481 PCHEPCKFSSYEKWKWNNFGQALPSODEGSKAFVNCDENSRLVSLTLNVTBDEG 540  
 QY 541 WYMGVKGQGHFYGETAAVYVAVERKAAGSRDVLAKADAPDEKVLDSGFREIENKAI 600  
 DB 541 WYMGVKGQGHFYGETAAVYVAVERKAAGSRDVLAKADAPDEKVLDSGFREIENKAI 600  
 QY 601 DPLFAFEKAVADTRQADGSRASVDSGSEEGGSSRALVSTLPGLVLAAGAVALGV 660  
 DB 601 DPLFAFEKAVADTRQADGSRASVDSGSEEGGSSRALVSTLPGLVLAAGAVALGV 660  
 QY 661 ARARHRKNDRVRSIRSYRTDISMSDFENSRFEGANDMGASSITQETSLGKEEYATTE 720  
 DB 661 ARARHRKNDRVRSIRSYRTDISMSDFENSRFEGANDMGASSITQETSLGKEEYATTE 720  
 QY 721 STTETKEPKKAKRSKEEAEMAKDVLLOSSTYAAAOQPOEA 764  
 DB 721 STTETKEPKKAKRSKEEAEMAKDVLLOSSTYAAAOQPOEA 764  
 RESULT 3  
 ID ABJ04350 standard; Protein: 764 AA.  
 XX ABJ04350;  
 AC 11-OCT-2002 (first entry)  
 XX  
 DE Human colon specific protein SEQ ID NO: 124.  
 XX  
 KW Human: colon specific gene; colon specific protein; colon cancer;  
 KW colorectal cancer; colon disease; cytostatic; gene therapy.  
 XX  
 OS Homo sapiens.  
 XX WO200242460-A2.  
 PN 30-MAY-2002.  
 XX  
 PD 21-NOV-2001; 2001WO-US43611.  
 PF 22-NOV-2000; 2000US-252505P.  
 PR (DIAD-) DIADEXUS INC.  
 PA Machina RA, Reclipon H, Plura J, Ghosh MG, Sun Y, Liu C;  
 PI WPI; 2002-583378/62.  
 DR Novel colon specific polypeptides and polynucleotides useful for  
 PT detecting, diagnosing, monitoring, treating, staging and predicting  
 PT cancers in humans having cancer and non-cancerous colon disease  
 PS Claim 11; Page 225-228; 228pp; English.  
 XX The present invention provides protein and coding sequences of human  
 CC colon specific genes and proteins. These can be used in the treatment of  
 CC colonic diseases, including colon and colorectal cancers. The present  
 CC sequence is a protein of the invention.

[illegible]

Key	Location/Qualifiers
13..45	putative immunoglobulin binding residues of domain I"
1..120	/label= domain_I
110..230	/label= domain_II
210..340	/label= domain_III
320..450	/label= domain_IV
440..550	/label= domain_V
550..606	/note= "external portions of domain VI"
550..627	/note= "external portions of domain VI"
625..660	/label= transmembrane_segment
650..746	/label= intracellular_portion
MO9621012-A1.	
11-JUL-1996.	
27-DEC-1995;	95WO-US16889.
04-MAY-1995;	95US-0434000.
30-DEC-1994;	94US-0367395.
(PLAN-) PLANT BIOTECHNOLOGY INC.	
(UNNE-) UNITED MEDICAL & DENTAL SCHOOLS GUYS.	
(PLAN-) PLANET BIOTECHNOLOGY INC.	
Hiatt AC, Lehner T, Ma JKC;	
WPI: 1996-333987/33.	
N-PSDB: AAT31288.	
Immunoglobulin and protection protein complex and its prodn. in plants - useful for passive immunisation against mucosal antigens, esp. against S. mutans and S. sorbitinus to prevent dental caries	
Disclosure: Pages 105-108; 152pp: English.	
The present sequence is the human poly-immunoglobulin (Ig) receptor, a portion of which corresp. to residues 1-627, pref. 1-606, or esp. residues 13-45, 1-120, 110-230, 210-340, 320-450, 440-550, 550-606 or 550-627 comprises a protection protein (PP). The Ig of the invention comprises a PP as above in association with an Ig derived heavy chain, having at least a portion of an antigen (Ag) binding domain. The PP protects the Ig in harsh mucosal, e.g. gastrointestinal, environments, therefore enhancing its effectiveness in passively immunising animals against mucosal pathogens. The Ag binding domain is specifically derived from the Guy's 13 antibody, and the Ig can be used to prevent dental caries by binding, e.g. Streptococcus mutans serotypes c, e and f, or S. sorbitinus serotypes d and g.	
Sequence 746 AA:	
Query Match 97.6%; Score 746; DB 17; Length 746;	
Best Local Similarity 100.0%; Pred. No. 0;	
Matches 746; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	

```

Db      1 KSPIFGPEEVNSVSGSVSTCYPTYSVNRHTRKWCROGARGCITLLISSEGYSSKY 60
QY      79 AGRANLTNPENGTFVYVNIQAOLSDODSGRYKCGLGINSRGLSPDVLSEYSGPGLNDTK 138
        |||||||
Db      61 AGRANLTNPENGTFFVYVNIQAOLSDODSGRYKCGLGINSRGLSPDVLSEYSGPGLNDTK 120
QY      139 VYTVDLGRTVTINCPRKTENAOKRKSLYKQIGLYPVLVIDSSGVNPNYTGRIQLDIOGT 198
        |||||||
Db      121 VYTVDLGRTVTINCPRKTENAOKRKSLYKQIGLYPVLVIDSSGVNPNYTGRIQLDIOGT 180
QY      199 GOLLEFSVINQLRLSDAGQYLCQAGDSDNSNKKNAOLYLKPEPELYEDLKGSVTFHCA 258
        |||||||
Db      181 GOLLEFSVINQLRLSDAGQYLCQAGDSDNSNKKNAOLYLKPEPELYEDLKGSVTFHCA 240
QY      259 LGPEVANVAKFLCRSGGSCNDVYVNTLGKRAPAFEGRIILNPODGSFSVYITLGRKE 318
        |||||||
Db      241 LGPEVANVAKFLCRSGGSCNDVYVNTLGKRAPAFEGRIILNPODGSFSVYITLGRKE 300
QY      319 DAGRYLCGAHSDGQLQEGSPIOAMOLFVNEESTIPRSPYVKGAGSSVAVLCPYNRKES 378
        |||||||
Db      301 DAGRYLCGAHSDGQLQEGSPIOAMOLFVNEESTIPRSPYVKGAGSSVAVLCPYNRKES 360
QY      379 KSIRKWCLEGAQNGRCPLVDSEGWKAOYEGRLSLLEPENGFTVILNDLTSRDAGF 438
        |||||||
Db      361 KSIRKWCLEGAQNGRCPLVDSEGWKAOYEGRLSLLEPENGFTVILNDLTSRDAGF 420
QY      439 YVCLTNGDTLMRTTVEIKIIEGEPNLKVPNGNTAVAGETLKYPCHPKFSSEYKXWCKM 498
        |||||||
Db      421 YVCLTNGDTLMRTTVEIKIIEGEPNLKVPNGNTAVAGETLKYPCHPKFSSEYKXWCKM 480
QY      499 NNTGCOALPDSODEGSPKAFVNCDENSRLVSLTLNLVTRADEGMYWCGVKQGHFYGETAAV 558
        |||||||
Db      481 NNTGCOALPDSODEGSPKAFVNCDENSRLVSLTLNLVTRADEGMYWCGVKQGHFYGETAAV 540
QY      559 YVAVEERKAGSRDVSILAKADAAPDEKVIDSGPRETENKAIDDPRLFAEKKAVADTRDA 618
        |||||||
Db      541 YVAVEERKAGSRDVSILAKADAAPDEKVIDSGPRETENKAIDDPRLFAEKKAVADTRDA 600
QY      619 DGSRAVSVDGSGSEEDGSSRALVSTLPLGLVLAAGAVAVGARARHRKRVDSIRSYR 678
        |||||||
Db      601 DGSRAVSVDGSGSEEDGSSRALVSTLPLGLVLAAGAVAVGARARHRKRVDSIRSYR 660
QY      679 TDISMSDFENSRFGANDNMGASITTOETSLGCKEEFVATTESTETETKPKKAKSSKEE 738
        |||||||
Db      661 TDISMSDFENSRFGANDNMGASITTOETSLGCKEEFVATTESTETETKPKKAKSSKEE 720
QY      739 AEMAYKDFLOSSTVAEAQDGPQEA 764
        |||||||
Db      721 AEMAYKDFLOSSTVAEAQDGPQEA 746

RESULT 5
AAM95601
ID      AAM95601 standard; Protein; 607 AA.
AC      AAM95601;
XX      08-JUN-1999 (first entry)
DT      XX
DE      Human secretory Immunoglobulin A component.
KW      Immunoglobulin A; secretory; component; IGA; human; treatment;
        prevention; HIV; AIDS; cold; flu; virus;
        human immunodeficiency virus; respiratory syncytial virus.
OS      Homo sapiens.
XX      XX
PN      WO9857993-A1.
XX      23-DEC-1998.
PD      XX
PF      10-JUN-1998; 98WO-US11975.
        XX

```

```

PR      19-JUN-1997; 97US-0050969.
XX      XX
PA      (REGC ) UNIV CALIFORNIA.
XX      XX
PI      Chintalacharuvu KR, Morrison SL;
XX      XX
DR      WPI: 1999-080950/07.
DR      N-PSDB: AAX07407.
XX      XX
PT      Producing secretory immunoglobulin in single cells - useful to
        produce commercial quantities of secretory immunoglobulin to prevent
        or treat infections
XX      XX
PS      Disclosure: Pages 22-24; 39pp; English.
XX      XX
CC      The sequence is that of the secretory component of human secretory
        immunoglobulin A (siga). It can be used as part of a method for
        the production of sig molecules. This method is useful for
        producing commercial quantities of sig (especially siga) to treat
        or prevent infections. In particular, siga produced by the method
        can be used to prevent or treat infections in mammals, birds or
        fish; especially systemic infections or infections at a mucosal
        surface. It is especially useful to prevent or treat infection
        with human immunodeficiency virus (HIV), respiratory syncytial
        virus, flu virus or cold virus. The method allows production of
        commercial quantities of sig molecules for therapeutic use, not
        previously possible: production using non-plant cells and a
        single cell type is more efficient than a previous multi-step
        process of fusing recombinant plant cells, and avoids alterations
        of the sig by plant cells. Siga molecules are more stable
        and resistant to proteolysis than previously used Iga molecules,
        and can be administered to prevent as well as to treat infections,
        unlike e.g. Igg and Igm molecules.
XX      XX
SQ      Sequence 607 AA.
        Query Match 79.5%; Score 607; DB 20; Length 607;
        Best Local Similarity 100.0%; Pred. No. 0;
        Matches 607; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY      1 MLFLVTLCLLAFPAISRTKSPFGEDEVNSVSGSVSTCYPTYSVNRHTRKWCROGA 60
        |||||||
Db      1 MLFLVTLCLLAFPAISRTKSPFGEDEVNSVSGSVSTCYPTYSVNRHTRKWCROGA 60
QY      61 RGCCTLLISSEGYVSSKXAGRANLTNPENGTFVYVNIQAOLSDODSGRYKCGLGINSRGLS 120
        |||||||
Db      61 RGCCTLLISSEGYVSSKXAGRANLTNPENGTFVYVNIQAOLSDODSGRYKCGLGINSRGLS 120
QY      121 FDVSLVSGPGLNDTKVYTVYDLGRTVTINCPRKTENAOKRKSLYKQIGLYPVLVIDSS 180
        |||||||
Db      121 FDVSLVSGPGLNDTKVYTVYDLGRTVTINCPRKTENAOKRKSLYKQIGLYPVLVIDSS 180
QY      181 GYVNPNTYGRIRLDIOGTGOLLEFSVINQLRLSDAGQYLCQAGDSDNSNKKNAOLYLK 240
        |||||||
Db      181 GYVNPNTYGRIRLDIOGTGOLLEFSVINQLRLSDAGQYLCQAGDSDNSNKKNAOLYLK 240
QY      241 EPELYEDLKGSVTFHCALEPEVANVAKFLCRSGGSCNDVYVNTLGKRAPAFEGRIILN 300
        |||||||
Db      241 EPELYEDLKGSVTFHCALEPEVANVAKFLCRSGGSCNDVYVNTLGKRAPAFEGRIILN 300
QY      301 PODKGSFSVYITLGRKEDAGRYLCGAHSDGQLQEGSPIOAMOLFVNEESTIPRSPYVVK 360
        |||||||
Db      301 PODKGSFSVYITLGRKEDAGRYLCGAHSDGQLQEGSPIOAMOLFVNEESTIPRSPYVVK 360
QY      361 GYAGSSVAVILCPYNRKESKSIKWKMLBGAQNGRCPLVDSEGWKAOYEGRLSLLEP 420
        |||||||
Db      361 GYAGSSVAVILCPYNRKESKSIKWKMLBGAQNGRCPLVDSEGWKAOYEGRLSLLEP 420
QY      421 NGTFVILNQLTRSDAGFYWCLTNGDTLMRTTVEIKIIEGEPNLKVPNGNTAVAGETLKY 480
        |||||||
Db      421 NGTFVILNQLTRSDAGFYWCLTNGDTLMRTTVEIKIIEGEPNLKVPNGNTAVAGETLKY 480
QY      481 PCHFPCKFSSEYKXWCKMNTGCOALPDSODEGSPKAFVNCDENSRLVSLTLNLVTRADEG 540
        |||||||

```



Db 481 PCHFPCKFSSEYKWKCKMNNNTGCOALPSODEGPKAAVNCDENSRLVSLTLNLTTRADG 540  
 Qy 541 WYMGVKGQHFYGETAAVYVAVEERKAAGSRDVSIAKADAAPEKVLDSGFREIENKAIO 600  
 Db 541 WYMGVKGQHFYGETAAVYVAVEERKAAGSRDVSIAKADAAPEKVLDSGFREIENKAIO 600  
 Qy 601 DPLRLFAE 607  
 Db 601 DPLRLFAE 607

## RESULT 6

AAV34099

ID AAV34099 standard; Protein: 607 AA.

AC AAV34099;

DT 20-DEC-1999 (first entry)

DE Partial amino acid sequence of plasmid pSHUSC.

KW Multimeric protein; immunoglobulin; receptor-ligand complex;

KW hetero-dimeric receptor; trimeric G protein; transgenic.

OS Synthetic.

PN MO9949024-A2.

PD 30-SEP-1999.

PE 24-MAR-1999; 99MO-US06506.

PR 25-MAR-1998; 98US-0079249.

PA (PLAN-) PLANET BIOTECHNOLOGY INC.

PI Wycoff KL, Jaiswal SK;

DR WPI: 1999-580446/49.

DR N-PSDB: AA222290.

PT Producing heterologous multimeric proteins in plants, transformed with  
 PT several plasmids expressing polypeptide components, particularly for  
 PT immunoglobulins

PS Example 1; Fig 8; 42pp: English.

CC The invention relates to a method for producing heterologous, multimeric  
 CC proteins in plant cells. The method comprises: (a) transforming the cells  
 CC with several naked plasmids each encoding some, but not all, of the  
 CC polypeptide components of the multimeric proteins, and together providing  
 CC all the polypeptide components; and (b) culturing the cells. The method  
 CC is used to produce biologically active multimeric proteins particularly  
 CC immunoglobulins, receptor-ligand complexes, homo- or hetero-dimeric  
 CC receptors, or trimeric G proteins. This method provides properly  
 CC associated and assembled multimeric proteins in a fast and efficient  
 CC process, without the need to cross plants expressing single component of  
 CC the protein. Transgenic plants containing adjacent and stably integrated  
 CC plasmids, and their progeny can also express the multimeric proteins. The  
 CC present sequence represents the partial amino acid sequence of the  
 CC plasmid pSHUSC.

SQ Sequence 607 AA;

Query Match 79.2%; Score 605; DB 20; Length 607;

Best Local Similarity 100.0%; Pred. No. 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 LFVLTCLLAVFAIPAISTKSPFGPEEVNSVSGNSITCYPPTSVNRHTRKRYWCROGARG 62  
 Db 3 LFVLTCLLAVFAIPAISTKSPFGPEEVNSVSGNSITCYPPTSVNRHTRKRYWCROGARG 62

Qy 63 GCITLISSEGVSSKTAGRANLTNFBENGTFVYVNTIQLSDDSDGRIKCGIGINSRLSFD 122  
 Db 63 GCITLISSEGVSSKTAGRANLTNFBENGTFVYVNTIQLSDDSDGRIKCGIGINSRLSFD 122  
 Qy 123 VSLSEVSGPCLLNDTFTVYVDLGRVTYINCPFTENAKRSLSYKQIGLPPVIVISSG 182  
 Db 123 VSLSEVSGPCLLNDTFTVYVDLGRVTYINCPFTENAKRSLSYKQIGLPPVIVISSG 182  
 Qy 183 VNPNTYGRIRLDIOGTGQLFSVYINQLRLSDAGQYLCAQADDNSNKKADQLVLPKP 242  
 Db 183 VNPNTYGRIRLDIOGTGQLFSVYINQLRLSDAGQYLCAQADDNSNKKADQLVLPKP 242  
 Qy 243 ELVYEDLRGVSFTFHCALGPEVAVNAKFLCROSSGECNDVYVNTLGRAPAFEGRIILLNP 302  
 Db 243 ELVYEDLRGVSFTFHCALGPEVAVNAKFLCROSSGECNDVYVNTLGRAPAFEGRIILLNP 302  
 Qy 303 DKDGSFVSVTGTLRKEDAGRYLCGANSDDQLDQCSPTIQAMQLFVNEESTIPRSPVYKGV 362  
 Db 303 DKDGSFVSVTGTLRKEDAGRYLCGANSDDQLDQCSPTIQAMQLFVNEESTIPRSPVYKGV 362  
 Qy 363 AGSSVAVLCPYNRKESKSIKYMCLMEGAQNGRCPLLVDSGMYKAQYEGRLSLLEPGNG 422  
 Db 363 AGSSVAVLCPYNRKESKSIKYMCLMEGAQNGRCPLLVDSGMYKAQYEGRLSLLEPGNG 422  
 Qy 423 TFTVILNQLTSRDAGFYWCLTNGDTLMRTTVEIKIIEGEPNLKYPGNVAVLGETLKYPC 482  
 Db 423 TFTVILNQLTSRDAGFYWCLTNGDTLMRTTVEIKIIEGEPNLKYPGNVAVLGETLKYPC 482  
 Qy 483 HPPCKFSSEYKWKCKMNNNTGCOALPSODEGPKAAVNCDENSRLVSLTLNLTTRADG 542  
 Db 483 HPPCKFSSEYKWKCKMNNNTGCOALPSODEGPKAAVNCDENSRLVSLTLNLTTRADG 542  
 Qy 543 WCGVKQHFYGETAAVYVAVEERKAAGSRDVSIAKADAAPEKVLDSGFREIENKAIO 602  
 Db 543 WCGVKQHFYGETAAVYVAVEERKAAGSRDVSIAKADAAPEKVLDSGFREIENKAIO 602  
 Qy 603 RLPFAE 607  
 Db 603 RLPFAE 607

## RESULT 7

AAM47867

ID AAM47867 standard; Protein: 607 AA.

AC AAM47867;

DT 22-FEB-2002 (first entry)

DE Human secretory component.

KW Human; immunoadhesin; intercellular adhesion molecule; ICAM-1;

KW human rhinovirus; immunoglobulin heavy chain; J chain; HRV; common cold;

KW transgenic plant.

OS Homo sapiens.

PN WO200183529-A2.

PD 08-NOV-2001.

PE 28-APR-2001; 2001MO-US13932.

PR 28-APR-2000; 2000US-200298P.

PA (PLAN-) PLANET BIOTECHNOLOGY INC.

PI Larrick JW, Wycoff KL;

DR WPI: 2002-041481/05.

DR N-PSDB: ABA05260.

PT Immunoadhesin for treating human rhinovirus infection comprises

PT chimeric intercellular adhesion molecule-1, and optionally a J chain  
 PT and secretory component in association  
 XX Examples: Fig 8; 138pp; English.  
 CC The invention relates to an immunoadhesin comprising:  
 CC (a) a chimeric intercellular adhesion molecule (ICAM)-1 comprising a  
 CC rhinovirus receptor protein linked to at least a portion of an  
 CC immunoglobulin heavy chain; and  
 CC (b) optionally a J chain and secretory component associated with the  
 CC chimeric ICAM-1 molecule. The immunoadhesin has plant-specific  
 CC glycosylation and virucide activity. The immunoadhesin is useful for  
 CC reducing infection by human rhinovirus (HRV) and hence the initiation or  
 CC spread of the common cold by HRV. The immunoadhesin binds to HRV and  
 CC reduces its infectivity, competing with cell surface ICAM-1 for binding  
 CC sites, interfering with virus entry or uncoating and directing premature  
 CC release of viral RNA and formation of empty capsids. Expression of the  
 CC immunoadhesin in plants would be tetrameric, rather than dimeric.  
 CC Immunoadhesin having multiple binding sites have a higher effective  
 CC affinity for the virus, thereby increasing the effectiveness of the  
 CC immunoadhesin. Association of secretory component and immunoglobulin J  
 CC chain increases the stability of the immunoadhesin in the mucosal  
 CC environment. Production is significantly less expensive in plants than in  
 CC animal cell culture and production in plants is safer for human use,  
 CC since plants are not known to harbor any animal viruses. The present  
 CC sequence is that of the human secretory component expressed from the  
 CC plasmid pShusc, of the invention.  
 XX  
 SO Sequence 607 AA:  
 Query Match 79.2%; Score 605; DB 23; Length 607;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 605; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Oy 3 LEVLTCLLAVPAISTSPKIPGPEEVNSVEGNSVITCYPTPTSYNRHTRKRCOGARG 62  
 Db 3 LEVLTCLLAVPAISTSPKIPGPEEVNSVEGNSVITCYPTPTSYNRHTRKRCOGARG 62  
 Oy 63 GCITLISSEGVSSKYGAGRANLTNFPENGTFVNNIAILOSDDSGRYKCGLGINSRGLSFD 122  
 Db 63 GCITLISSEGVSSKYGAGRANLTNFPENGTFVNNIAILOSDDSGRYKCGLGINSRGLSFD 122  
 Oy 123 VSELEVSOGPLLNTKTYVTVDLGRVTNCFEFTENAKRKSLYKQIGLYVLYVYDSSGY 182  
 Db 123 VSELEVSOGPLLNTKTYVTVDLGRVTNCFEFTENAKRKSLYKQIGLYVLYVYDSSGY 182  
 Oy 183 VNPVYTGRIIDIGTGGLFSVYINOLRLSDAGCYLCQAGDDSNKKNADLOYLKEP 242  
 Db 183 VNPVYTGRIIDIGTGGLFSVYINOLRLSDAGCYLCQAGDDSNKKNADLOYLKEP 242  
 Oy 243 ELVYEDLRGVTFFHACALGPEVANYAKFLCRQSSGENDVYVNTLGRAPAFEGRIILLNPQ 302  
 Db 243 ELVYEDLRGVTFFHACALGPEVANYAKFLCRQSSGENDVYVNTLGRAPAFEGRIILLNPQ 302  
 Oy 303 DKDSFSVYITGLKKEADAGRTLCGAHSDGQLQEGSPIQANOLFYNEESTIRSPYVYKGV 362  
 Db 303 DKDSFSVYITGLKKEADAGRTLCGAHSDGQLQEGSPIQANOLFYNEESTIRSPYVYKGV 362  
 Oy 363 AGSSVAVLCPYNRKESKIKYKWLMEGAONGRCPLVDSEGMVKAQYGRSLLEEPGNG 422  
 Db 363 AGSSVAVLCPYNRKESKIKYKWLMEGAONGRCPLVDSEGMVKAQYGRSLLEEPGNG 422  
 Oy 423 TFFVILNOLTRDAGFYWCLTNGDTLMRTTVEIKIIEGEPNLKYPGNTAVLGETLKYPC 482  
 Db 423 TFFVILNOLTRDAGFYWCLTNGDTLMRTTVEIKIIEGEPNLKYPGNTAVLGETLKYPC 482  
 Oy 483 HFPCFSSYEYKWKNNNTGCOALPSODEGSKAFVNCDENSRLVSLTLNLYTRADGMY 542  
 Db 483 HFPCFSSYEYKWKNNNTGCOALPSODEGSKAFVNCDENSRLVSLTLNLYTRADGMY 542  
 Oy 543 MCGVKOGHFYGETAAVYAVEERKAAGSDVSLAKADAAPDEKVLDSGFREIENKAIDDP 602  
 Db 543 MCGVKOGHFYGETAAVYAVEERKAAGSDVSLAKADAAPDEKVLDSGFREIENKAIDDP 602

Oy 603 RLEAE 607  
 Db 603 RLEAE 607  
 RESULT 8  
 ID AAY73981 standard; Protein; 272 AA.  
 XX AAY73981;  
 AC AAY73981;  
 XX 14-MAR-2000 (first entry)  
 DE Human prostate tumor EST fragment derived protein #168.  
 XX Pancreas: tumor; EST; expressed sequence tag; human; cytostatic;  
 KW treatment.  
 XX Homo sapiens.  
 OS DE19820190-A1.  
 XX DE19820190-A1.  
 XX 04-NOV-1999.  
 PD 28-APR-1998; 98DE-1020190.  
 XX 28-APR-1998; 98DE-1020190.  
 XX 28-APR-1998; 98DE-1020190.  
 XX (META-) METAGEN GES GENOMFORSCHUNG MBH.  
 XX Rosenthal A, Specht T, Hinzmann B, Schmitt A, Pilarsky C, Dahl E;  
 DR MPI: 1999-621386/54.  
 XX N-PSDB; AA252913.  
 PT New human nucleic acid sequences from pancreatic tumors, and related  
 PS proteins  
 XX  
 PS Claim 23; Page 379; 502pp; German.  
 CC This invention describes novel polypeptides and their encoding nucleic  
 CC acids derived from human pancreatic tumor tissue which have cytostatic  
 CC activity. The sequences are also useful in producing pharmaceutical  
 CC compositions for treatment of pancreatic tumors. AAY73814-174252  
 CC represent protein fragments encoded by the human pancreatic tumor cDNA  
 CC library derived expressed sequence tag (EST) sequences represented in  
 CC AA252858-253014.  
 XX  
 SO Sequence 272 AA:  
 Query Match 33.2%; Score 254; DB 20; Length 272;  
 Best Local Similarity 100.0%; Pred. No. 27e-251;  
 Matches 254; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Oy 511 EGPSKAFVNCDENSRLVSLTLNLYTRADGMYKGVGOGHFYGETAAVYAVEERKAAGS 570  
 Db 19 EGPSKAFVNCDENSRLVSLTLNLYTRADGMYKGVGOGHFYGETAAVYAVEERKAAGS 78  
 Oy 571 RDVSLAKADAAPDEKVLDSGFREIENKAIDDPRLFABEKAAVDROADDSRAVDGSS 630  
 Db 79 RDVSLAKADAAPDEKVLDSGFREIENKAIDDPRLFABEKAAVDROADDSRAVDGSS 138  
 Oy 631 EEOGSSRALVSTLVPGLVAVGAVAVARARRKRVDSIRSYRTDISMSDFENSR 690  
 Db 139 EEOGSSRALVSTLVPGLVAVGAVAVARARRKRVDSIRSYRTDISMSDFENSR 198  
 Oy 691 EFGANDNMGASSITQETSLGCKEEFVAFTTSTETTKPKKAKRSKEBAEMAYKDFLLS 750  
 Db 199 EFGANDNMGASSITQETSLGCKEEFVAFTTSTETTKPKKAKRSKEBAEMAYKDFLLS 258  
 Oy 751 STVAEADODGPOEA 764  
 |||||||||||||||

Db 259 STVAEAKQDQPEA 272

# RESULT 9

ABP55307  
ID ABP55307 standard; Protein: 243 AA.

AC ABP55307;

DT 28-JAN-2003 (first entry)

DE Human polyimmunoglobulin receptor (pIgR) amino acid sequence.

KM Transseptinial transport; membrane bound vesicle; virion; liposome;  
KM envelope; capsid; transmembrane domain; gene therapy; immunostimulant;  
KM cytosolic; haemostatic; neuroprotective; antirheumatic; antitubercular;  
KM anticancer; antibacterial; anti-HIV; hepatotropic; virucide; exocytosis;  
KM antiinflammatory; apical endocytosis; basolateral endocytosis; ADA-SCID;  
KM transcytosis; monogenic disease; ADA deficiency; cystic fibrosis; ALS;  
KM X-linked severe combined immunodeficiency; Haemophilia B; cancer; HIV;  
KM chronic granulomatous disease; coronary artery disease; viral infection;  
KM amyotrophic lateral sclerosis; rheumatoid arthritis; hepatitis; Herpes;  
KM pathogenic disorder; human immunodeficiency virus; bacterial infection;  
KM tuberculosis; Chlamydia; gastroenteritis; ulcer; pIgR;  
KM polyimmunoglobulin receptor.

OS Homo sapiens.

PN WO200283840-A2.

PD 24-OCT-2002.

PF 03-APR-2002; 2002WO-US10647.

PR 03-APR-2001; 2001US-281275P.

PA (ARIZ-) ARIZEXE PHARM INC.

PI Sheridan PL, Houston LL;

DR WPI; 2003-046923/04.

PT Fusion protein which confers the ability to penetrate epithelial cell  
PT layer and to undergo paracellular transport, has a transseptinial  
PT delivery element and a transmembrane domain from different proteins  
PS Disclosure: Fig 2B: 160pp; English.

CC The present invention describes a fusion protein (I) comprising a  
CC transseptinial delivery element (TDE) from a first protein and a  
CC transmembrane domain from a second protein, or comprising TDE and a  
CC viral sequence that confers the ability to be associated with or  
CC incorporated into an envelope or capsid protein of a virus. (I) has  
CC immunostimulant, cytosolic, haemostatic, neuroprotective, antirheumatic,  
CC antitubercular, anticancer, antibacterial, anti-HIV, hepatotropic,  
CC virucide and antiinflammatory activities, and can be used in gene  
CC therapy. (I) confers the ability to undergo apical endocytosis,  
CC basolateral endocytosis, apical or basolateral exocytosis, apical to  
CC basolateral transcytosis and basolateral to apical transcytosis. Diseases  
CC treatable by gene therapy include monogenic diseases such as X-linked  
CC severe combined immunodeficiency, ADA deficiency (ADA-SCID), cystic  
CC fibrosis, Haemophilia B, chronic granulomatous disease, cancers such as  
CC ovarian cancer, other diseases such as coronary artery disease,  
CC amyotrophic lateral sclerosis (ALS), rheumatoid arthritis, pathogenic  
CC disorders, including human immunodeficiency virus (HIV), viral  
CC infections, hepatitis, non-specific bacterial infection, tuberculosis,  
CC Herpes, Chlamydia and gastroenteritis. The present sequence  
CC represents a polyimmunoglobulin receptor (pIgR) amino acid sequence  
CC which is given in the exemplification of the present invention.

CC Sequence 243 AA;

Query Match

31.8%; Score 243; DB 24; Length 243;

Best Local Similarity 100.0%; Pred. No. 4.6e-240; Indels 0; Gaps 0;  
Matches 243; Conservative 0; Mismatches 0;

QY 492 EKYMCKNNNTGCOALPDSODEGPKAFVNCDENSRLVSLTNLYTRADGMYGCVQGHF 551  
DB 1 EKYMCKNNNTGCOALPDSODEGPKAFVNCDENSRLVSLTNLYTRADGMYGCVQGHF 60  
QY 552 YGETAAVYVAVEERKAAGSRDYSKADAAPEDEKVLDSGFREIENKAIDPRLFAEKAV 611  
DB 61 YGETAAVYVAVEERKAAGSRDYSKADAAPEDEKVLDSGFREIENKAIDPRLFAEKAV 120  
QY 612 ADTRODADSRASVDSGSSSEEGGSSRALVSTPLGLVLAAGAVALGARRRHRKNDR 671  
DB 121 ADTRODADSRASVDSGSSSEEGGSSRALVSTPLGLVLAAGAVALGARRRHRKNDR 180  
QY 672 VSTRSRRTDMSDFNSRPFANDMGASTTOETSLGKEEFPATTESTETKEPKKA 731  
DB 181 VSTRSRRTDMSDFNSRPFANDMGASTTOETSLGKEEFPATTESTETKEPKKA 240  
QY 732 KRS 734  
DB 241 KRS 243

## RESULT 10

ABP55315  
ID ABP55315 standard; Protein: 243 AA.

AC ABP55315;

DT 28-JAN-2003 (first entry)

DE Human polyimmunoglobulin receptor (pIgR) amino acid sequence.

KM Transseptinial transport; membrane bound vesicle; virion; liposome;  
KM envelope; capsid; transmembrane domain; gene therapy; immunostimulant;  
KM cytosolic; haemostatic; neuroprotective; antirheumatic; antitubercular;  
KM anticancer; antibacterial; anti-HIV; hepatotropic; virucide; exocytosis;  
KM antiinflammatory; apical endocytosis; basolateral endocytosis; ADA-SCID;  
KM transcytosis; monogenic disease; ADA deficiency; cystic fibrosis; ALS;  
KM X-linked severe combined immunodeficiency; Haemophilia B; cancer; HIV;  
KM chronic granulomatous disease; coronary artery disease; viral infection;  
KM amyotrophic lateral sclerosis; rheumatoid arthritis; hepatitis; Herpes;  
KM pathogenic disorder; human immunodeficiency virus; bacterial infection;  
KM tuberculosis; Chlamydia; gastroenteritis; ulcer; pIgR;  
KM polyimmunoglobulin receptor.

OS Homo sapiens.

PN WO200283840-A2.

PD 24-OCT-2002.

PF 03-APR-2002; 2002WO-US10647.

PR 03-APR-2001; 2001US-281275P.

PA (ARIZ-) ARIZEXE PHARM INC.

PI Sheridan PL, Houston LL;

DR WPI; 2003-046923/04.

PT Fusion protein which confers the ability to penetrate epithelial cell  
PT layer and to undergo paracellular transport, has a transseptinial  
PT delivery element and a transmembrane domain from different proteins  
PS Disclosure: Fig 2D: 160pp; English.

CC The present invention describes a fusion protein (I) comprising a  
CC transseptinial delivery element (TDE) from a first protein and a  
CC transmembrane domain from a second protein, or comprising TDE and a  
CC viral sequence that confers the ability to be associated with or

CC Incorporated into an envelope or capsid protein of a virus. (1) has  
 CC immunostimulant, cytostatic, haemostatic, neuroprotective, antirheumatic,  
 CC antitachytic, antitumor, antibacterial, anti-HIV, hepatotropic,  
 CC virucide and antiinflammatory activities, and can be used in gene  
 CC therapy. (1) confers the ability to undergo apical endocytosis.  
 CC basolateral endocytosis, apical or basolateral exocytosis, apical to  
 CC basolateral transcytosis and basolateral to apical transcytosis. Diseases  
 CC treatable by gene therapy include monogenic diseases such as X-linked  
 CC severe combined immunodeficiency, ADA deficiency (ADA-SCID), cystic  
 CC fibrosis, Haemophilia B, chronic granulomatous disease, cancers such as  
 CC ovarian cancer, other diseases such as coronary artery disease,  
 CC amyotrophic lateral sclerosis (ALS), rheumatoid arthritis, pathogenic  
 CC disorders, including human immunodeficiency virus (HIV), viral  
 CC infections, hepatitis, non-specific bacterial infection, tuberculosis,  
 CC Herpes, Chlamydia and gastrointestinal ulcer. The present sequence  
 CC represents a polyimmunoglobulin receptor (PIgR) amino acid sequence  
 CC which is given in the exemplification of the present invention.

XX Sequence 243 AA:

Query Match 31.8%; Score 243; DB 24; Length 243;  
 Best Local Similarity 100.0%; Pred. No. 4,6e-240;  
 Matches 243; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 492 EKWKCKNNNTGQALPSQDGPSPKAFVNCBENSLVLTNLVTRADEGMYCCKGKH 551  
 |||||||  
 Db 1 EKYCKNNNTGQALPSQDGPSPKAFVNCBENSLVLTNLVTRADEGMYCCKGKH 60  
 OY 552 YGTAAYVAVEERKKAAGSRDVSIAKADAPDEKVLDSGFEIENKAIDPRLFAEKKAV 611  
 |||||||  
 Db 61 YGTAAYVAVEERKKAAGSRDVSIAKADAPDEKVLDSGFEIENKAIDPRLFAEKKAV 120  
 OY 612 ADTRDQADGSRASVDSGSSSEEGSSRALVSTVPLDLYAVGAVAVGARARRKKAVDR 671  
 |||||||  
 Db 121 ADTRDQADGSRASVDSGSSSEEGSSRALVSTVPLDLYAVGAVAVGARARRKKAVDR 180  
 OY 672 VTSRSTYRTDIMSDFENSRFGANDNMGASSTOETSLGKKEEVAATTESTTEKPKKA 731  
 |||||||  
 Db 181 VTSRSTYRTDIMSDFENSRFGANDNMGASSTOETSLGKKEEVAATTESTTEKPKKA 240  
 OY 732 KRS 734  
 ||||  
 Db 241 KRS 243

RESULT 11

ABP5311 standard; Protein: 94 AA.

XX ABP5311:  
 AC 28-JAN-2003 (first entry)  
 DT Human polyimmunoglobulin receptor (PIgR) stalk region.  
 DE  
 XX  
 XX Trans epithelial transport; membrane bound vesicle; virion; liposome;  
 KW envelope; capsid; transmembrane domain; gene therapy; immunostimulant;  
 KW cyostatic; haemostatic; neuroprotective; antirheumatic; antitachytic;  
 KW antitumor; antibacterial; anti-HIV; hepatotropic; virucide; exocytosis;  
 KW antiinflammatory; apical endocytosis; basolateral endocytosis; ADA-SCID;  
 KW transcytosis; monogenic disease; ADA deficiency; cystic fibrosis; ALS;  
 KW X-linked severe combined immunodeficiency; Haemophilia B; cancer; HIV;  
 KW chronic granulomatous disease; coronary artery disease; viral infection;  
 KW amyotrophic lateral sclerosis; rheumatoid arthritis; hepatitis; Herpes;  
 KW pathogenic disorder; human immunodeficiency virus; bacterial infection;  
 KW tuberculosis; Chlamydia; gastrointestinal ulcer; PIgR;  
 KW polyimmunoglobulin receptor.  
 XX  
 XX Homo sapiens.  
 OS  
 XX  
 XX MO200283840-A2.  
 PN  
 XX  
 PD 24-OCT-2002.

XX 03-APR-2002; 2002WO-US10647.  
 PF  
 XX 03-APR-2001; 2001US-281275P.  
 PR  
 XX (ARI2-) ARIEKE PHARM INC.  
 PA  
 XX Sheridan PL, Houston LL;  
 PI  
 XX WPI; 2003-046923/04.  
 DR  
 XX

PT Fusion protein which confers the ability to penetrate epithelial cell  
 layer and to undergo paracellular transport, has a trans epithelial  
 PT delivery element and a transmembrane domain from different proteins  
 PS Disclosure; Fig 2C; 160pp; English.

XX The present invention describes a fusion protein (1) comprising a  
 CC trans epithelial delivery element (TDE) from a first protein and a  
 CC transmembrane domain from a second protein, or comprising TDE and a  
 CC viral sequence that confers the ability to be associated with or  
 CC incorporated into an envelope or capsid protein of a virus. (1) has  
 CC immunostimulant, cytostatic, haemostatic, neuroprotective, antirheumatic,  
 CC antitachytic, antitumor, antibacterial, anti-HIV, hepatotropic,  
 CC virucide and antiinflammatory activities, and can be used in gene  
 CC therapy. (1) confers the ability to undergo apical endocytosis,  
 CC basolateral endocytosis, apical or basolateral exocytosis, apical to  
 CC basolateral transcytosis and basolateral to apical transcytosis. Diseases  
 CC treatable by gene therapy include monogenic diseases such as X-linked  
 CC severe combined immunodeficiency, ADA deficiency (ADA-SCID), cystic  
 CC fibrosis, Haemophilia B, chronic granulomatous disease, cancers such as  
 CC ovarian cancer, other diseases such as coronary artery disease,  
 CC amyotrophic lateral sclerosis (ALS), rheumatoid arthritis, pathogenic  
 CC disorders, including human immunodeficiency virus (HIV), viral  
 CC infections, hepatitis, non-specific bacterial infection, tuberculosis,  
 CC Herpes, Chlamydia and gastrointestinal ulcer. The present sequence  
 CC represents a polyimmunoglobulin receptor (PIgR) amino acid sequence  
 CC which is given in the exemplification of the present invention.

SO Sequence 94 AA:

Query Match 12.3%; Score 94; DB 24; Length 94;  
 Best Local Similarity 100.0%; Pred. No. 1.1e-87;  
 Matches 94; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 545 GYKOGHYGETAAYVAVEERKKAAGSRDVSIAKADAPDEKVLDSGFEIENKAIDPRL 604  
 |||||||  
 Db 1 GYKOGHYGETAAYVAVEERKKAAGSRDVSIAKADAPDEKVLDSGFEIENKAIDPRL 60  
 OY 605 FAEKAVADTRDQADGSRASVDSGSSSEEGGSSR 638  
 |||||||  
 Db 61 FAEKAVADTRDQADGSRASVDSGSSSEEGGSSR 94

RESULT 12

AAG65712 standard; protein: 90 AA.

XX AAG65712:  
 AC 07-JAN-2002 (first entry)  
 DT Human polymeric immunoglobulin receptor (PIgR) fragment.  
 DE  
 XX  
 XX Human polymeric immunoglobulin receptor; PIgR; ligand; therapeutic;  
 KW carcinoma diagnosis; veterinary; human.  
 KW  
 XX  
 XX Homo sapiens.  
 OS  
 XX  
 XX MO200172846-A2.  
 PN  
 XX  
 PD 04-OCT-2001.

PF 26-MAR-2001: 2001MO-US09699.  
 XX  
 XX 27-MAR-2000: 2000US-192197P.  
 PR 27-MAR-2000: 2000US-192198P.  
 XX  
 PA (REGC ) UNIV CALIFORNIA.  
 XX  
 PI Mostov KE, Chapin SJ, Richman-Elisenstat J;  
 XX  
 DR WPI; 2001-611619/70.  
 XX  
 PT New ligands binding to a specific region of a polymetric immunoglobulin  
 PT receptor, useful for transporting therapeutic or diagnostic  
 PT compositions into or across cells expressing pigR e.g. in drug delivery  
 PS  
 XX  
 PS Disclosure: Fig 3: 102pp: English.  
 CC The invention provides ligands that bind specifically to a region of an  
 CC animal cell polymetric immunoglobulin receptor (pigR). The pigR cleaves  
 CC to produce a stalk region remaining attached to the cell and a secretory  
 CC component existing in the organ of interest in several forms. The ligands  
 CC do not bind to the stalk or the most abundant form of the secretory  
 CC component present in the organ under physiological conditions. The  
 CC ligands are useful for transporting therapeutic or diagnostic  
 CC compositions into or across cells expressing pigR, useful to introduce  
 CC or transport ligands such as antibodies and/or to deliver biologically  
 CC active components such as proteins, nucleic acids or detectable labels.  
 CC They are used to deliver therapeutic compositions to mucosal surfaces  
 CC such as the gastro-intestinal tract, respiratory system etc. in humans.  
 CC They are also useful to label cells expressing pigR, e.g. to distinguish  
 CC epithelial cells from a mixed cell population in pathology studies or to  
 CC aid in carcinoma diagnosis (since pigR expression is reduced in  
 CC carcinomas relative to normal epithelium). They can also be used to  
 CC deliver veterinary compositions, especially in mammals such as farm,  
 CC domestic or wild mammals or birds e.g. birds reared for human  
 CC consumption. The present sequence represents a human pigR fragment.  
 CC  
 XX  
 SO Sequence 90 AA:  
 Query Match 11.8%; Score 90; DB 22; Length 90;  
 Best Local Similarity 100.0%; Pred. No. 1.3e-83;  
 Matches 90; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 549 GHFYGETAAVYVVEERKAGSRDYS LAKADAAPEKVLDSGFREIENKAIDPRLFAEE 608  
 DB 1 GHFYGETAAVYVVEERKAGSRDYS LAKADAAPEKVLDSGFREIENKAIDPRLFAEE 60  
 QY 609 KAVADTRDQADGSRASVDSGSSSEEGGSSR 638  
 DB 61 KAVADTRDQADGSRASVDSGSSSEEGGSSR 90  
 RESULT 13  
 AAM43098  
 ID AAM43098 standard; peptide: 61 AA.  
 AC AAM43098;  
 XX  
 DT 04-JUN-1998 (first entry)  
 DE Polymetric immunoglobulin receptor (pigR) stalk sequence 1.  
 XX  
 KW Polymetric immunoglobulin receptor; pigR; stalk; epithelial cell;  
 KW ligand; antibody; target; binding; mammalian.  
 XX  
 OS Mammalia.  
 XX  
 PN WO9746588-A1.  
 XX  
 PD 11-DEC-1997.  
 PR 14-MAY-1997; 97WO-US07944.  
 PF

XX  
 PR 04-JUN-1996; 96US-0018958.  
 XX  
 XX  
 PA (REGC ) UNIV CALIFORNIA.  
 XX  
 PI Mostov KE, Richman-Elisenstat J;  
 XX  
 DR WPI; 1998-042123/04.  
 XX  
 PT ligand that binds the stalk of a cell's polymetric immunoglobulin  
 PT receptor - useful to target to, into or across mammalian epithelial  
 PT cell biologically active component, e.g. nucleic acid, protein,  
 PT lipid, carbohydrate, etc  
 PS  
 XX  
 PS Claim 28; Page 37; 42pp: English.  
 CC This peptide sequence represents the stalk of the polymetric  
 CC immunoglobulin receptor (pigR) to which a ligand can bind to. The stalk  
 CC is the extracellular component of the pigR that is bound to the cell  
 CC following cleavage of the secretory component of the pigR. The stalk is  
 CC present regardless of whether the secretory component segment is cleaved  
 CC or uncleaved from pigR. A ligand, preferably a humanised antibody or a  
 CC recombinant single chain variable region fragment can specifically bind  
 CC to the stalk of a pigR of a cell under physiological conditions, but not  
 CC to the secretory component of pigR. Such a ligand can be introduced into  
 CC a cell expressing a pigR by attaching to the stalk of the pigR. The  
 CC ligand can be used to target to, into or across the apical or basolateral  
 CC surface of a mammalian epithelial cell, a biologically active component  
 CC selected from a nucleic acid (preferably encoding the wild type cystic  
 CC fibrosis transmembrane conductance regulator), protein, radioisotope,  
 CC lipid or carbohydrate. The biologically active composition can also be  
 CC selected from a group consisting of anti-inflammatory, antisense  
 CC oligonucleotides, antibiotics or anti-infectives.  
 CC  
 XX  
 SO Sequence 61 AA:  
 Query Match 8.0%; Score 61; DB 19; Length 61;  
 Best Local Similarity 100.0%; Pred. No. 4.9e-54;  
 Matches 61; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 578 ADAAPDEKVLDSGFREIENKAIDPRLFAEEKAVADTRDQADGSRASVDSGSSSEEGGSS 637  
 DB 1 ADAAPDEKVLDSGFREIENKAIDPRLFAEEKAVADTRDQADGSRASVDSGSSSEEGGSS 60  
 QY 638 R 638  
 DB 61 R 61  
 RESULT 14  
 AAM43099  
 ID AAM43099 standard; peptide: 61 AA.  
 AC AAM43099;  
 XX  
 DT 04-JUN-1998 (first entry)  
 DE Polymetric immunoglobulin receptor (pigR) stalk sequence 2.  
 XX  
 KW Polymetric immunoglobulin receptor; pigR; stalk; epithelial cell;  
 KW ligand; antibody; target; binding; mammalian.  
 XX  
 OS Mammalia.  
 XX  
 PN WO9746588-A1.  
 XX  
 PD 11-DEC-1997.  
 PR 14-MAY-1997; 97WO-US07944.  
 PF 04-JUN-1996; 96US-0018958.  
 PR (REGC ) UNIV CALIFORNIA.  
 PA

```

XX  Moslov KE, Richman-Eisenstat J;
PI
XX
XX  WPI: 1998-042123/04.
DR
XX
XX  Ligand that binds the stalk of a cell's polymeric immunoglobulin
PT receptor - useful to target to, into or across mammalian epithelial
PT cell biologically active component, e.g. nucleic acid, protein,
PS lipid, carbohydrate, etc
XX
PS  Claim 28: Page 37; 42pp; English.
XX
XX  This peptide sequence represents the stalk of the polymeric
CC immunoglobulin receptor (PIGR) to which a ligand can bind to. The stalk
CC is the extracellular component of the PIGR that is bound to the cell
CC following cleavage of the secretory component of the PIGR. The stalk is
CC present regardless of whether the secretory component segment is cleaved
CC or uncleaved from PIGR. A ligand, preferably a humanised antibody or a
CC recombinant single chain variable region fragment can specifically bind
CC to the stalk of a PIGR of a cell under physiological conditions, but not
CC to the secretory component of PIGR. Such a ligand can be introduced into
CC a cell expressing a PIGR by attaching to the stalk of the PIGR. The
CC ligand can be used to target to, into or across the apical or basolateral
CC surface of a mammalian epithelial cell, a biologically active component
CC selected from a nucleic acid (preferably encoding the wild type cystic
CC fibrosis transmembrane conductance regulator), protein, radioisotope,
CC lipid or carbohydrate. The biologically active composition can also be
CC selected from a group consisting of anti-inflammatory, antisense
CC oligonucleotides, antibiotics or anti-infectives.
XX
XX  Sequence 61 AA:
SQ
Query Match 8.0%; Score 61; DB 19; Length 61;
Best Local Similarity 100.0%; Pred. No. 4.9e-54;
Matches 61; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Oy 578 ADAAPDEKVDLSGFRLEENKAIDPRLFAEKKAVADTRDQADGSRASVDSGSSSEEGGSS 637
Db 1 ADAAPDEKVDLSGFRLEENKAIDPRLFAEKKAVADTRDQADGSRASVDSGSSSEEGGSS 60
Oy 638 R 638
Db 61 R 61

```

RESULT 15  
AAM99073  
ID AAM99073 standard; Protein: 60 AA.  
XX  
XX AAM99073:  
AC  
XX  
XX 13-MAY-1999 (first entry)  
DT  
XX  
XX Immunoglobulin like protein PolyIgRv4.  
DE  
XX  
XX Human; PIGR-1; immunoglobulin; rheumatoid arthritis; psoriasis;  
KW multiple sclerosis; systemic lupus erythematosus; diagnosis; receptor;  
KW inflammatory bowel disease.  
XX  
XX OS Unidentified.  
XX  
XX EP897981-A1.  
PN  
XX  
XX 24-FEB-1999.  
PD  
XX  
XX 11-AUG-1998; 98EP-0306403.  
PE  
XX  
XX 22-OCT-1997; 97US-0955937.  
PR  
XX 19-AUG-1997; 97US-0056152.  
PR  
XX (SMIK ) SMITHKLINE BEECHAM.  
PA  
XX  
XX Hurtle MR, Sweet RW, Truneh A, Wu S;  
PI

```

XX  WPI: 1999-134644/12.
DR
XX
XX  New receptor (PIGR-1) polypeptide and polynucleotide - useful as
PT diagnostic reagents and for prevention and treatment of multiple
PT sclerosis, inflammatory bowel disease and psoriasis
XX
XX  Example 2: Page 14; 28pp; English.
PS
XX
XX  The present invention describes a new receptor polypeptide designated
CC PIGR-1, which is a member of the immunoglobulin (Ig) superfamily. PIGR-1
CC proteins and polynucleotides are useful for diagnosing susceptibility to
CC diseases by detecting mutations in the PIGR-1 gene, and can diagnose
CC diseases associated with PIGR-1 protein imbalance by determining PIGR-1
CC protein expression levels. PIGR-1 proteins can be used to screen for
CC agonists and antagonists by measuring the binding to protein, and
CC observing the protein function. These can be used in treatment to
CC activate (agonist) or inhibit (antagonist) PIGR-1 activity, in addition
CC to direct administration of antisense sequences to prevent expression, or
CC PIGR-1 polynucleotides to treat conditions associated with a lack of
CC PIGR-1 protein. Gene therapy may also be used to affect endogenous PIGR-1
CC protein expression. PIGR-1 antibodies are useful for inducing an immune
CC response to immunise and prevent disease, and for isolating PIGR-1
CC clones or purifying the polypeptides by affinity chromatography. PIGR-1
CC proteins can be administered directly or as a vaccine to inoculate
CC against disease. Diseases diagnosed, prevented and treated include:
CC rheumatoid arthritis, multiple sclerosis, psoriasis, systemic lupus
CC erythematosus and inflammatory bowel disease. The PIGR-1 protein is also
CC useful for mapping the gene to a chromosome, allowing gene inheritance
CC to be studied through linkage analysis. The present sequence represents
CC an immunoglobulin like protein from the present invention.
XX
XX  Sequence 60 AA:
SQ
Query Match 7.9%; Score 60; DB 20; Length 60;
Best Local Similarity 100.0%; Pred. No. 5.1e-53;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Oy 353 PRSPVTYKGVAGSSVAVLCPPYNRKESKSIKYWCLMEGAONRCPLVDSEGVKAYQYEGR 412
Db 1 PRSPVTYKGVAGSSVAVLCPPYNRKESKSIKYWCLMEGAONRCPLVDSEGVKAYQYEGR 60

```

Search completed: September 9, 2003, 23:17:14  
Job time : 86 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 9, 2003, 23:15:47 ; Search time 32 Seconds  
(without alignments)  
1010.172 Million cell updates/sec

Title: US-09-818-247-1

Perfect score: 764  
Sequence: 1 MLFLVLTCLAVPAISTKS.....DFLLQSTVAAEADGPQEA 764

Scoring table: OLIGO  
Gapop 60.0 , Gapext 60.0

Searched: 328717 seqs, 42310858 residues

Word size : 0

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database :

Issued Patents\_Aa: \*  
1: /cgn2\_6/ptodata/1/1aa/5A.COMB.pep: \*  
2: /cgn2\_6/ptodata/1/1aa/5B.COMB.pep: \*  
3: /cgn2\_6/ptodata/1/1aa/5A.COMB.pep: \*  
4: /cgn2\_6/ptodata/1/1aa/5B.COMB.pep: \*  
5: /cgn2\_6/ptodata/1/1aa/PCTUS.COMB.pep: \*  
6: /cgn2\_6/ptodata/1/1aa/backfile1.pep: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	746	97.6	746	3	US-08-434-000A-4
2	746	97.6	746	4	US-09-312-157-4
3	607	79.5	608	4	US-09-095-385-4
4	61	8.0	61	3	US-08-856-383-10
5	61	8.0	61	3	US-08-856-383-11
6	61	8.0	61	4	US-09-475-088-10
7	61	8.0	61	4	US-09-475-088-11
8	60	7.9	60	3	US-08-955-937A-6
9	60	7.9	60	3	US-09-300-985-6
10	53	6.9	109	3	US-08-961-564A-9
11	43	5.6	43	3	US-08-955-937A-11
12	43	5.6	43	3	US-09-300-985-11
13	31	4.1	31	3	US-08-856-383-4
14	31	4.1	31	4	US-09-475-088-4
15	29	3.8	46	3	US-08-955-937A-10
16	29	3.8	46	3	US-09-300-985-10
17	27	3.5	57	3	US-08-955-937A-5
18	27	3.5	57	3	US-09-300-985-5
19	21	2.7	757	3	US-08-434-000A-6
20	21	2.7	757	4	US-09-312-157-6
21	21	2.5	40	3	US-08-856-383-6
22	22	2.5	40	4	US-09-475-088-6
23	17	2.2	769	3	US-08-434-000A-10
24	17	2.2	769	4	US-09-312-157-10
25	16	2.1	16	3	US-08-782-480-45
26	16	2.1	16	3	US-08-954-211-45
27	16	2.1	16	4	US-09-005-167A-45

28	16	2.1	16	4	US-09-176-741B-45	Sequence 45, Appl
29	16	2.1	771	3	US-08-434-000A-8	Sequence 8, Appl
30	16	2.1	771	4	US-09-312-157-8	Sequence 8, Appl
31	13	1.7	773	3	US-08-434-000A-2	Sequence 2, Appl
32	13	1.7	773	4	US-09-312-157-2	Sequence 2, Appl
33	12	1.6	16	2	US-08-656-906-1	Sequence 1, Appl
34	12	1.6	16	3	US-09-217-847-1	Sequence 1, Appl
35	9	1.2	624	2	US-08-642-406A-22	Sequence 22, Appl
36	9	1.2	624	4	US-09-199-534-22	Sequence 22, Appl
37	9	1.2	624	4	US-09-199-534-22	Sequence 22, Appl
38	8	1.0	60	3	US-08-961-564A-6	Sequence 6, Appl
39	8	1.0	417	4	US-09-252-991B-27213	Sequence 27213, A
40	8	1.0	777	2	US-08-874-678-3	Sequence 3, Appl
41	8	1.0	777	3	US-08-643-839-3	Sequence 3, Appl
42	8	1.0	777	4	US-09-348-886-3	Sequence 3, Appl
43	8	1.0	802	4	US-09-173-151A-33	Sequence 33, Appl
44	8	1.0	1298	1	US-08-222-616-33	Sequence 33, Appl
45	8	1.0	1298	1	US-08-340-011-2	Sequence 2, Appl

#### ALIGNMENTS

RESULT 1  
US-08-434-000A-4  
Sequence 4, Application US/08434000A  
Patent No. 6046037  
GENERAL INFORMATION:  
APPLICANT: ANDREW C. HIATT, JULIAN  
APPLICANT: K.-C. MA, THOMAS LEHNER  
TITLE OF INVENTION: IMMUNOGLOBULINS CONTAINING PROTECTION  
TITLE OF INVENTION: PROTEINS IN PLANTS AND THEIR USES  
NUMBER OF SEQUENCES: 19  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Lyon & Lyon  
STREET: 633 West Fifth Street  
CITY: Los Angeles  
STATE: California  
COUNTRY: U.S.A.  
ZIP: 90071  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5" Diskette, 1.44 MB  
MEDIUM TYPE: storage  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: IBM P.C. DOS 5.0  
SOFTWARE: Word Perfect 5.1  
CURRENT APPLICATION DATA: US/08/434,000A  
APPLICATION NUMBER: US/08/434,000A  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA: Including application  
PRIOR APPLICATION DATA: described below:  
APPLICATION NUMBER: 08/367,395  
ATTORNEY/AGENT INFORMATION:  
NAME: Guise, Jeffrey W.  
REGISTRATION NUMBER: 34,613  
REFERENCE/DOCKET NUMBER: 212/127  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (619) 552-8400  
TELEFAX: (619) 552-0159  
TELEX: 67-3510  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 746 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
? ? ? ? ? Human Polymunoglobulin Receptor  
US-08-434-000A-4

Query Match 97.6%; Score 746; DB 3; Length 746;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 746; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

19 KSPIFGPEEVSNSVSGSITCYPTSVNRHTRKRYWCROGARGCITLISSEGVSSKY 78  
1 KSPIFGPEEVSNSVSGSITCYPTSVNRHTRKRYWCROGARGCITLISSEGVSSKY 60  
79 AGRANLTNPENGTFFVNIQAQLSDSGRYKCGLGINSRGLSPDVSLEVSQGGLLNDK 138  
61 AGRANLTNPENGTFFVNIQAQLSDSGRYKCGLGINSRGLSPDVSLEVSQGGLLNDK 120  
139 VYTVDLGRVTINCPKTEAOKRSKLYKQIGLYPVLVYDSSGVNPNYGRIRLDIOGT 198  
121 VYTVDLGRVTINCPKTEAOKRSKLYKQIGLYPVLVYDSSGVNPNYGRIRLDIOGT 180  
199 GOLFFSVINQLRLSDAGQYLCQAGDSSNKKNNADLOVLPPELYEDLRGSVTFHCA 258  
181 GOLFFSVINQLRLSDAGQYLCQAGDSSNKKNNADLOVLPPELYEDLRGSVTFHCA 240  
259 LGPEVANNAKFLCROSSGENDVYVNTLGKRAFAFEGRIILNPQDKGSESVITGLRKE 318  
241 LGPEVANNAKFLCROSSGENDVYVNTLGKRAFAFEGRIILNPQDKGSESVITGLRKE 300  
319 DAGRYLCGASHDQLOEGSPIQAMOLFVNEESTIPSPYVKGAGSSAVVLCPPYRKES 378  
301 DAGRYLCGASHDQLOEGSPIQAMOLFVNEESTIPSPYVKGAGSSAVVLCPPYRKES 360  
379 KSIKYWCLMEGAONGRCPLLYDSEGWKAQYEGRLSLLEPGNGTFTVILNQLTSRDAGF 438  
361 KSIKYWCLMEGAONGRCPLLYDSEGWKAQYEGRLSLLEPGNGTFTVILNQLTSRDAGF 420  
439 YWCLTNGDTLMRTTVEIKIIEGEPNLKVPGNVAVLGETLKVCHFPCKSSYEKYWKW 498  
421 YWCLTNGDTLMRTTVEIKIIEGEPNLKVPGNVAVLGETLKVCHFPCKSSYEKYWKW 480  
499 NNTGCOALPSQDEGSPSAFVNCDENSRLVSLTLNLTVRADGWTWCVKQGHFYGETAAV 558  
481 NNTGCOALPSQDEGSPSAFVNCDENSRLVSLTLNLTVRADGWTWCVKQGHFYGETAAV 540  
559 YVAVEERKAASRDVSLAKADAAPEDEVLDSGFREIENKAIODPRLEAEKAAVADTRQA 618  
541 YVAVEERKAASRDVSLAKADAAPEDEVLDSGFREIENKAIODPRLEAEKAAVADTRQA 600  
619 DGSRAVDSGSSSEBOGSSRALVSTVPLGLVLAAGVAVARARRKRVDRYSISYR 678  
601 DGSRAVDSGSSSEBOGSSRALVSTVPLGLVLAAGVAVARARRKRVDRYSISYR 660  
679 TDISMDFENSREPGANDNGASSITQETSLGKEEVAATTESTETKEPKAKRSSKEE 738  
661 TDISMDFENSREPGANDNGASSITQETSLGKEEVAATTESTETKEPKAKRSSKEE 720  
739 AEMAYKDFLLQSSSTVAEAODGPOEA 764  
721 AEMAYKDFLLQSSSTVAEAODGPOEA 746

RESULT 2  
US-09-312-157-4  
Sequence 4, Application US/09312157  
Patent No. 6303341  
GENERAL INFORMATION:  
APPLICANT: ANDREW C. HIATT, JULIAN  
K.-C. MA, THOMAS LEHNER  
TITLE OF INVENTION: IMMUNOGLOBULINS CONTAINING PROTECTION  
PROTEINS IN PLANTS AND THEIR USES  
NUMBER OF SEQUENCES: 19  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Lyon & Lyon  
STREET: 633 West Fifth Street  
Suite 4700  
CITY: Los Angeles

STATE: California  
COUNTRY: U.S.A.  
ZIP: 90071  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5" Diskette, 1.44 MB  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: IBM P.C. DOS 5.0  
SOFTWARE: Word Perfect 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/312,157  
FILING DATE: 14-May-1999  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/434,000  
FILING DATE: <Unknown>  
ATTORNEY/AGENT INFORMATION:  
NAME: Guise, Jeffrey W.  
REGISTRATION NUMBER: 34,613  
REFERENCE/DOCKET NUMBER: 212/127  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (619) 552-8400  
TELEFAX: (619) 552-0159  
TELEX: 67-351  
SEQUENCE LISTING  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 746 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
DESCRIPTION: Human Polymunoglobulin Receptor  
SEQUENCE DESCRIPTION: SEQ ID NO: 4  
US-09-312-157-4

Query Match 97.6%; Score 746; DB 4; Length 746;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 746; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
19 KSPIFGPEEVSNSVSGSITCYPTSVNRHTRKRYWCROGARGCITLISSEGVSSKY 78  
1 KSPIFGPEEVSNSVSGSITCYPTSVNRHTRKRYWCROGARGCITLISSEGVSSKY 60  
79 AGRANLTNPENGTFFVNIQAQLSDSGRYKCGLGINSRGLSPDVSLEVSQGGLLNDK 138  
61 AGRANLTNPENGTFFVNIQAQLSDSGRYKCGLGINSRGLSPDVSLEVSQGGLLNDK 120  
139 VYTVDLGRVTINCPKTEAOKRSKLYKQIGLYPVLVYDSSGVNPNYGRIRLDIOGT 198  
121 VYTVDLGRVTINCPKTEAOKRSKLYKQIGLYPVLVYDSSGVNPNYGRIRLDIOGT 180  
199 GOLFFSVINQLRLSDAGQYLCQAGDSSNKKNNADLOVLPPELYEDLRGSVTFHCA 258  
181 GOLFFSVINQLRLSDAGQYLCQAGDSSNKKNNADLOVLPPELYEDLRGSVTFHCA 240  
259 LGPEVANNAKFLCROSSGENDVYVNTLGKRAFAFEGRIILNPQDKGSESVITGLRKE 318  
241 LGPEVANNAKFLCROSSGENDVYVNTLGKRAFAFEGRIILNPQDKGSESVITGLRKE 300  
319 DAGRYLCGASHDQLOEGSPIQAMOLFVNEESTIPSPYVKGAGSSAVVLCPPYRKES 378  
301 DAGRYLCGASHDQLOEGSPIQAMOLFVNEESTIPSPYVKGAGSSAVVLCPPYRKES 360  
379 KSIKYWCLMEGAONGRCPLLYDSEGWKAQYEGRLSLLEPGNGTFTVILNQLTSRDAGF 438  
361 KSIKYWCLMEGAONGRCPLLYDSEGWKAQYEGRLSLLEPGNGTFTVILNQLTSRDAGF 420  
439 YWCLTNGDTLMRTTVEIKIIEGEPNLKVPGNVAVLGETLKVCHFPCKSSYEKYWKW 498  
421 YWCLTNGDTLMRTTVEIKIIEGEPNLKVPGNVAVLGETLKVCHFPCKSSYEKYWKW 480  
499 NNTGCOALPSQDEGSPSAFVNCDENSRLVSLTLNLTVRADGWTWCVKQGHFYGETAAV 558